Theory on light scattering by metal nano-structures, by way of the Green tensor method

Departamento de Física de la Materia Condensada
Instituto de Ciencia de Materiales de Aragón
CSIC-Universidad de Zaragoza

DOCTORAL THESIS

Theory on light scattering by metal nano-structures, *by way of the Green tensor method.*

Giovanni Brucoli

Thesis advisor:

Luis Martín Moreno

Zaragoza, Mayo 2010

Ai miei genitori.

Contents

Acknowledgements

I would like to thank my thesis advisor Luis for all I have learned in these years as a PhD student with him. Also thanks to F.J. García and Andrey Evlyukhin for their help with my thesis. Thanks to I. P. Radko, S. I. Bozhevolnyi, Alexey, F.de León it was nice working with you!
Thanks to all my friends in the Zaragoza condensed matter department for cheering me up all the times. Thanks to the rest of my friends in the Facultad de Ciencias, INA and the rest of them in zaragoza. Also thanks to all my brothers and sisters back home in Bari for understanding me just like I am, for being on my side and for our summer-time sessions.

Thanks to my friends from all over the "Global Village", from New Zealand and Singapore all the way to Santo Domingo and Argentina (passing through London and Rome) for showing me what I can be (we are FAMILY).

Finally I feel compelled to acknowledge that I belong to a lucky (privileged) minority of the human family that made it this far, while the overwhelming majority of human beings on the Planet struggle for access to basic resources, *including education*. The argument that this is so because of some transcendental flaw of human nature is a myth. It's time for us as a people to start making some changes. Let's change the way we eat, let's change the way we live and let's change the way we treat each other.

Introduction

This thesis reports a theoretical investigation of the interaction between light and metal nano-metric structures on a thick metal slab. When light is confined in a sub-wavelength structure, the electric field inside such structure is attained by solving Maxwell's equations self-consistently. This can be done by means of the Green Tensor Approach (GTA). Not only does the GTA provide the field within a sub-wavelength nano-metric structure, it also allows for an analytic study of the light scattered by the nano-structure into the surrounding space, far from the structure. Nonetheless the implementation of the Green tensor approach is rather complex, mathematically, especially if the considered scattering centers interact with a metal slab, rather than being placed in vacuum. In fact accounting for the presence of a back-ground metal slab, demands performing Sommerfeld's integrals.

The first part of this thesis is devoted to reporting the formalism of the electromagnetic Green tensor in a complete manner, for all of the considered nano-optical systems. Chapters 1-3 detail the mathematical foundations of the numerical codes by means of which the results of this thesis were attained.

In Chapter 1 we define and derive the electromagnetic Green tensor for a semi-space back-ground. Our derivation is an organic elaboration of many disjointed pieces of standard material. These have been originally assembled and linked by detailed derivations, which are either briefly sketched in the literature or not mentioned.

In Chapter 2 we describe our implementation of the procedure to obtain numerical convergence of Sommerfeld's integrals, describing the interaction of fields with a metal plane. Part of the numerical procedure had appeared previously in the literature and so we refer to it as the standard integration technique. However after describing thoroughly our implementation of the standard technique, we also report on our modification to the standard integration technique, which speeds up computation in those cases in which surface plasmon polaritons constitute the main mechanism of light transport on the metal plane.

In Chapter 3 we derive the asymptotic expressions of the Green tensor that should be used to calculate, analytically, the Poynting vector energy flux scat-

tered by bi-dimensional nano-structures. These expressions in the considered case of bi-dimensional systems are not found in the literature, to the best of our knowledge. Furthermore we have developed a simplified formalism that deduces the fields amplitude, in the far-zone, from scalar products rather than through more tedious tensorial operations.
The second part of the thesis is devoted to reporting the results we have attained for several nano-optical systems.
In Chapter 4 we report our investigation of a device to launch a *unidirectional* stream of surface plasmon polariton, based on photonic band-gap effects. The device can achieve *local light coupling* into surface plasmon polariton modes. This work is original and, besides having potential technological application, it provides evidence of our theory through experimentally replicated results.
In Chapter 5 we study the effect on the propagation of a surface plasmon caused by an impedance barrier. This represents a metal wire of rectangular cross-section in a thick conducting film. By varying the in-plane direction of propagation of the surface plasmon with respect to the normal to the impedance barrier, we find a surface plasmon analog of the Brewster angle. At such angle the incident surface plasmon is not reflected by the impedance barrier.
In Chapter 6 we begin a comparative study of the surface plasmon scattering by bi-dimensional protrusions (ridges) and indentations (grooves). Subwavelength protrusions and indentations of equal shape present different scattering coefficients when their height and width are comparable. In this case, a protrusion scatters plasmons like a vertical point-dipole on a plane, while an indentation scatters like a horizontal point-dipole on a plane. We corroborate that, as previously presented in the literature with approximate methods, long and shallow asymmetrically-shaped surface defects have very similar scattering. Moreover we provide a first principles explanation for such property. In the transition from short shallow scatterers to long shallow scatterers the radiation is explained in terms of interference between a vertical and a horizontal dipole. The results attained numerically are exact and accounted for with analytical models.
In Chapter 7 we extend the comparative study of surface plasmon scattering by ridges and grooves. This time the width of the defects is fixed, while their height is varied. Both individual and arrays of defects are considered, mainly in the optical regime. The width of the defects is fixed, while their height is varied. It is shown that protrusions mainly reflect the incident plasmons in the optical range. Indentations, however, mainly radiate the incident plasmon out of plane. An indentation produces maximum reflection and out-of-plane radiation at the same wavelength, when its interaction with the incident surface plasmon is resonant. Protrusions, in general, exhibit maximum reflection and radiation at different wavelengths. Shallow arrays of either defects produce a

photonic band-gap, whose spectral width can be broadened by increasing the defects height or depth. At wavelengths inside the band-gap ridge arrays reflect SPPs better than groove arrays, while groove arrays radiate SPPs better than ridge arrays.

In Chapter 8 the scattering of an incident SPP by a three-dimensional nano-particle is analyzed. SPP extinction spectra are calculated for gold cubic particles of various sizes placed in the vicinity of a flat gold surface, as the distance of the particle from the flat gold surface is varied. The results are compared with an analytical model in which the interaction of the nano-particle with the plane is approximated by that of a point-dipole.

Chapter 1

Green's Function: Definition and derivation

1.1 Electromagnetic Green Tensor

The Green function is a mathematical tool to find an integral solution to a linear partial inhomogeneous differential equation[1]. As we shall see, physically it can be interpreted as a space-time propagator. In our treatment, we shall yield steady-state solution of Maxwell's equations, assuming on all quantities a harmonic time dependence of the kind $\mathbf{A}(\mathbf{r},t) = \mathbf{A}(\mathbf{r})e^{-i\omega t}$. If needed, the response for wave packets can be obtained by a Fourier transform in the frequency domain consisting of monochromatic waves. Therefore Green's function represents here the space propagator.
From first principles we know that Maxwell's equations are an abstraction resulting from reassigning the mechanical concept of force to the structure of space itself, by means of the concept of Field. Accordingly, the Electric field is defined as the electric force per unit charge. Green's function goes one step further, as it is the electric force on test point charge at $\mathbf{r}$, exerted by a *test point source* at $\mathbf{r}'$.
For a scalar wave, given a unit source at a point r', Green's function quantifies its electromagnetic interaction with a point of space r, by weighting it over distance and phase difference. More generally in 3D a unit source $\mathbf{r}'$ is typified by a point-dipole and each orthogonal component of the dipole independently generates a vector field at an arbitrary point $\mathbf{r}'$. Therefore we can condense the emission of the three orthogonal point dipoles on $\mathbf{r}$, into a tensor derived from Green's function. This is Green's tensor [2] and is used to include the changes in direction of the source-field through space.
In this chapter we are going to highlight a brief and convenient derivation of

the Green Tensor, with a derivation that is a synthesis of different sources [1–13]. We are going to work in the cgs system wherein the speed of light is $c = 1$ and in free-space the frequency coincides with the wavevector $\omega = g$ where $g = \sqrt{k_x^2 + k_y^2 + k_z^2}$.
Our final aim is to arrive at the Green tensor necessary for treating the scattering of Electric fields by surface defects in a plane surface. To do this we employ The Green Tensor Approach. This starts by considering an electric field $\mathbf{E}_b(\mathbf{r})$ in an arbitrary unperturbed background with dielectric constant ε_b, which we shall regard as 'background' field. The background field satisfies the homogenous wave equation:

$$\nabla \times \nabla \times \mathbf{E}_b(\mathbf{r}) - \varepsilon_b\, g^2 \;\; \mathbf{E}_b(\mathbf{r}) \;\; = \;\; 0. \tag{1.1}$$

The backgrounds we are going to treat are either free space or two bounding air-metal semi-spaces (eg. space is filled with air from $z > 0$ and field with a metal from $z < 0$). Nonetheless we do not need to specify the background right now, because the following derivation has general validity for an arbitrary background.
Now if the background generating the field of Eq.(1.1), is disrupted to accommodate a scatterer localized in a region V where the dielectric constant suffers a variation $\Delta\varepsilon = \varepsilon - \varepsilon_b$, the net field $\mathbf{E}(\mathbf{r})$ in a region V satisfies a new homogenous wave equation:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \varepsilon g^2 \;\; \mathbf{E}(\mathbf{r}) \;\; = \;\; 0 \tag{1.2}$$

this is a source-less equation restricted to the volume V with permittivity ε. If we use the superposition principle we can define a source density equal to the induced polarization [14]: $\mathbf{P}(\mathbf{r}) = \Delta\varepsilon\, \mathbf{E}(\mathbf{r})$ to transform Eq.(1.2) from source-less to an equivalent driven equation in the unperturbed background [11] :

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \varepsilon_b\, g^2\, \mathbf{E}(\mathbf{r}) = g^2 \mathbf{P}(\mathbf{r}). \tag{1.3}$$

The source term is defined in the volume V as: $\mathbf{\Theta}(\mathbf{r}) = ig\mathbf{J}(\mathbf{r}) = g^2\mathbf{P}(\mathbf{r}) = g^2\Delta\varepsilon\, \mathbf{E}(\mathbf{r})$. Notice that whether $\mathbf{\Theta}(\mathbf{r})$ is capacitevily induced or generated by passing a current through volume V, the solution scheme this equation leads to will be the same mathematically. Solving Eq.(1.3) requires at least a homogenous solution plus a particular solution. We have one homogeneous solution which is just the physical illumination on the defect Eq.(1.1). We need to add a particular solution to the inhomogeneous and possibly the rest of the complementary solutions to Eq.(1.1). This takes us to to the definition of the Green's Tensor.
We define the Green Tensor $\hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}')$ as the solution to:

$$\mathbf{\Theta}(\mathbf{r}) = \int_V d\mathbf{r}' \;\; \left[\nabla \times \nabla \times \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{\Theta}(\mathbf{r}') - \varepsilon_b g^2 \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{\Theta}(\mathbf{r}')\right] \tag{1.4}$$

where $\nabla = \nabla_r$ throughout.
If we find such tensor then we have constructed a particular solution $\mathbf{E}_s(\mathbf{r})$ of the field:

$$\mathbf{E}_s(\mathbf{r}) = \int_V d\mathbf{r}' \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \boldsymbol{\Theta}(\mathbf{r}') \tag{1.5}$$

as taking the differential operators out of the integral in Eq.(1.4) and using Eq.(1.5) we get:

$$\boldsymbol{\Theta}(\mathbf{r}) = \nabla \times \nabla \times \mathbf{E}_s(\mathbf{r}) - \varepsilon_b \, g^2 \, \mathbf{E}_s(\mathbf{r}), \tag{1.6}$$

which is a solution to Eq.(1.3).
Hence to construct the electric field in an inhomogeneity we need to find the tensor $\hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}')$. From its definition in Eq.(1.4) we can get the differential equation for $\hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}')$ if we attain an integral representation of the driving field $\boldsymbol{\Theta}(\mathbf{r})$ using the concept of localized unit point-source $\delta(\mathbf{r} - \mathbf{r}')$:

$$\boldsymbol{\Theta}(\mathbf{r}) = \int_V d\mathbf{r}' \; \hat{\mathbf{1}} \cdot \boldsymbol{\Theta}(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') \tag{1.7}$$

if we now plug this equation into Eq.(1.4) and we end up with:

$$\nabla \times \nabla \times \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') - g^2 \, \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') = \hat{\mathbf{1}} \delta(\mathbf{r} - \mathbf{r}') \tag{1.8}$$

This tells us that the Green tensor represents the fields produced by any of the three components of a unit source placed at $\mathbf{r}'$ on the point $\mathbf{r}$. Hence Eq.(1.5) is a mathematical trick for finding the field $\mathbf{E}(\mathbf{r})$ generated by an extended source $\boldsymbol{\Theta}(\mathbf{r})$, by breaking it down into a superposition of elementary fields $\hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}')$ generated by infinitesimal impulse drives $\hat{\mathbf{1}}\delta(\mathbf{r} - \mathbf{r}')$. This is useful because the Green operator satisfies the same differential structure but has the advantage to be valid for an a point-like unit source, subject to the boundary conditions of the background space. The net field $\mathbf{E}(\mathbf{r})$ is actually determined by the effective geometry of the finite-sized source distribution $\boldsymbol{\Theta}(\mathbf{r})$ throughout V. As Eq.(1.6) holds simultaneously for all points in V, intuitively speaking, the field at any point $\mathbf{r}$ in V, must be correlated or interconnected to the field at any other point $\mathbf{r}'$ in V in a self-consistent manner. In fact Eq.(1.4) implies:

$$\mathbf{E}(\mathbf{r}) = \int_V d\mathbf{r}' \; \left[\nabla \times \nabla \times \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{E}(\mathbf{r}') - g^2 \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{E}(\mathbf{r}') \right] \tag{1.9}$$

When we solve Eq.(1.8) and apply the background boundary conditions to its solutions, we will attain how a point source $\mathbf{r}'$ radiates the electric field to any other point $\mathbf{r}$ in V. Hence we will build the electric field of Eq.(1.5) in the finite-sized source V as a superposition of the individual contribution from every point source in V.

1.2 Free-space Green's tensor spatial and spectral representations

The Green's Tensor of Eq.(1.8) is the general solution for an arbitrary background. In order to solve the equation we have to specify the system boundary conditions. As a start, let us get the solution in free-space $\hat{\mathbf{G}}_0(\mathbf{r},\mathbf{r}')$.
Notice that in any homogenous isotropic medium, the Green tensor must have translational symmetry $\hat{\mathbf{G}}_0(\mathbf{r},\mathbf{r}') = \hat{\mathbf{G}}_0(\mathbf{r}-\mathbf{r}') = \hat{\mathbf{G}}_0(\mathbf{R})$, where $\mathbf{R} = \mathbf{r}-\mathbf{r}'$.
Using the identity $\nabla\times\nabla = -\nabla^2 + \nabla\nabla$ we end up with:

$$\nabla^2\hat{\mathbf{G}}_0(\mathbf{R}) + g^2\hat{\mathbf{G}}_0(\mathbf{R}) = -\left(\hat{\mathbf{1}} + \frac{\nabla\nabla}{g^2}\right)\delta(\mathbf{R}) \tag{1.10}$$

where we take the divergence of Eq.(1.8), as prescribed in [2], to get

$$\nabla\cdot\hat{\mathbf{G}}_0(\mathbf{R}) = -\frac{\nabla\delta(\mathbf{R})}{g^2} \tag{1.11}$$

since $\nabla\cdot(f\hat{\mathbf{1}}) = \nabla f$.
Equation (1.11) is a shortcut that is less rigorous mathematically than the precise and well known treatment (eg. Ref.[15]). Nevertheless equation (1.11) does not alter the correct result, when it is cautionary integrated over a surface, as we shall substantiate in more detail in the next sections. Ultimately, we will get to focus on the physics of the GT in a more direct way.
The free space GT has both a spatial representation, as a function of $\mathbf{R}$ only, and a spectral representation obtained by Fourier transforming the spatial representation. In the case of the free-space or homogenous medium background the GT can be entirely derived from its spatial representation[2, 15]. Here we construct the GT from its spectral representation because this will be useful when finding the GT in a semi-space back-ground, where no spatial representation can be derived analytically. There are other similar derivations in the spectral domain[16], but the following derivation is our own work.
We express the vectors in a short-hand notation as: $\mathbf{R} = (\mathbf{R}_\parallel, Z)$, $\mathbf{k} = (\mathbf{k}_\parallel, k_z)$ where the subscript $\parallel$ indicates the component of the vector tangential to the XY plane. The choice of axes is arbitrary in free-space but shall result convenient later, when we extend the problem to a semi-space, because then we set metal surface at the XY plane. In this approach, while $k_\parallel$ is varied, k_z is automatically fixed to $k_z = (g^2 - k_\parallel^2)^{1/2}$. Notice that k_z is a complex square root, in our case we need to choose it such as $\Im\{k_z\} > 0$, to satisfy the Radiation Boundary Condition[2].
Now, the GT Fourier transform is defined by the expansion in plane waves of its spatial representation:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \hat{\mathbf{G}}_0(\mathbf{R}_\parallel, Z) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_\parallel\, e^{i\mathbf{k}_\parallel\cdot\mathbf{R}_\parallel}\ \hat{\mathbf{G}}_0(\mathbf{k}_\parallel, Z) \tag{1.12}$$

where we are using the following designation: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel} = \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y$ Having defined $\hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z)$ we can Fourier transform Eq.1.10 by integrating it over all space $\int_{-\infty}^{\infty}\ \int_{-\infty}^{\infty} d\mathbf{R}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}$, yielding:

$$\frac{d^2\hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z)}{dz^2} - k_{\parallel}^2\ \hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z) + g^2\ \hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z) = -\left(\hat{\mathbf{1}} - \frac{\mathbf{kk}}{g^2}\right)\delta(Z) \quad (1.13)$$

$$\frac{d^2\hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z)}{dz^2} + k_z^2\ \hat{\mathbf{G}}_0(\mathbf{k}_{\parallel}, Z) = -\left(\hat{\mathbf{1}} - \frac{\mathbf{kk}}{g^2}\right)\delta(Z) \quad (1.14)$$

Remember the gradient operator in the position space is represented by the wavevector in the reciprocal space $\nabla \to \pm i\mathbf{k}$.
Also notice that in the last equation $\mathbf{kk}$ is a tensor (a dyadic) and the notation consisting of two vectors with no dot in the middle, represents the outer product. This should not be confused with the scalar product of two vectors which we shall always indicate with a dot in-between the two vectors.
As we are about to show, the Green Tensor can be derived from its scalar Green function, similarly to how the electric field can be derived from its potential. In order to do this we need to focus on a convenient expression for the identity matrix. This can be obtained, respectively, in terms the wavevector $\mathbf{k}$ and of the electromagnetic eigenmodes of vacuum. The latter are divided into $\mathbf{k}_p$ p-polarization(TM) unit wave-modes and $\mathbf{k}_p$ s-polarization(TE) unit wave-modes. The set we need to consider is:

$$\mathbf{k}/g = \frac{1}{g}\left(k_x\mathbf{u}_x + k_y\mathbf{u}_y + k_z\mathbf{u}_z\right) \quad (1.15)$$

$$\mathbf{k}_p^{\pm} = \frac{k_z}{g\,k_{\parallel}}\left(\ k_x\mathbf{u}_x + k_y\mathbf{u}_y \mp \frac{k_{\parallel}}{k_z}\mathbf{u}_z\ \right) \quad (1.16)$$

$$\mathbf{k}_s = \frac{1}{k_{\parallel}}(k_y\mathbf{u}_x - k_x\mathbf{u}_y) \quad (1.17)$$

where $k_{\parallel} = \sqrt{k_x^2 + k_y^2}$. Notice that $\mathbf{k}_p^+ \neq \mathbf{k}_p^-$ and that $\mathbf{k}_s^+ = \mathbf{k}_s^- = \mathbf{k}_s$. This emerges by the constraint $\nabla\cdot\mathbf{e}_p^{\pm}(\mathbf{r}) = \nabla\cdot\mathbf{e}_s^{\pm}(\mathbf{r}) = 0$, where :

$$\mathbf{e}_p^{\pm}(\mathbf{r}) = e^{\mathbf{i}\mathbf{k}_{\parallel}\cdot\mathbf{r}_{\parallel}} e^{\pm\mathbf{i}k_z z}\ \mathbf{k}_p^{\pm} \quad (1.18)$$

is a p-polarized plane wave, and :

$$\mathbf{e}_s^{\pm}(\mathbf{r}) = e^{\mathbf{i}\mathbf{k}_{\parallel}\cdot\mathbf{r}_{\parallel}} e^{\pm\mathbf{i}k_z z}\ \mathbf{k}_s \quad (1.19)$$

is an s-polarized plane wave.

The unit wavevector and the vectorial part of these waves form a com-

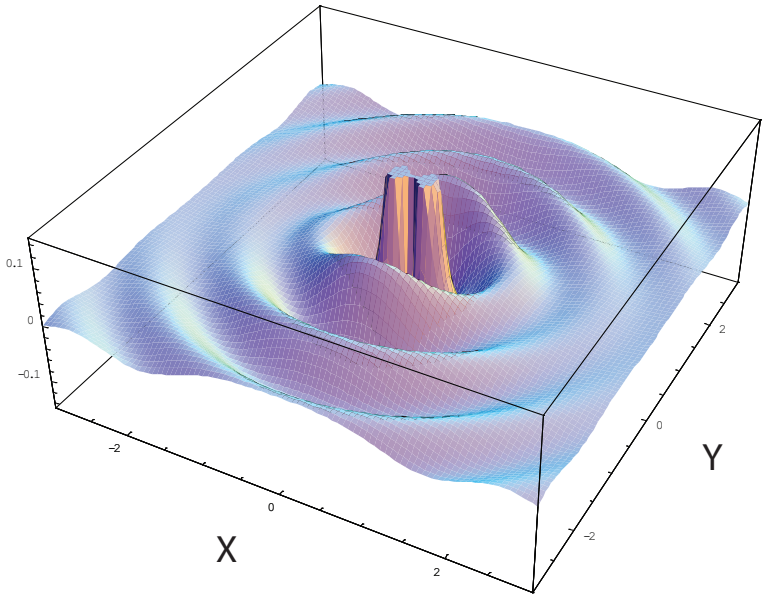

Figure 1.1: The Green tensor component $\Re\{G_{xx}(X,Y,0)\}$ in real space, as a function of the X and Y.

plete orthonormal basis. In fact changing to spherical coordinates $k_x = g\cos\phi\sin\theta$, $k_y = g\sin\phi\sin\theta$, $k_z = g\cos\theta$ we see that they are equivalent to the local frame of reference basis for the $\mathbf{k}$-space: $\mathbf{k}/g = \mathbf{u}_{k_r}$; $\mathbf{k}_p^{\pm} = \mathbf{u}_{k_\theta}$ and $\mathbf{k}_s = \mathbf{u}_{k\phi}$.
Thus expressing the identity matrix in the spherical coordinates basis of the reciprocal space yields:

$$\hat{\mathbf{1}} = \left(\mathbf{kk}/g^2 + \mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s\right) \tag{1.20}$$

Using this identity Eq.(1.14) can be written as:

$$\frac{d^2\hat{\mathbf{G}}_0(\mathbf{k}_\parallel, Z)}{dz^2} + k_z^2\ \hat{\mathbf{G}}_0(\mathbf{k}_\parallel, Z) = -\left(\mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s\right)\delta(Z)\,. \tag{1.21}$$

If we now assume there exists a scalar function $G_0(\mathbf{k}_\parallel, Z)$ such that: $\hat{\mathbf{G}}_0(\mathbf{k}_\parallel, Z) = G_0(\mathbf{k}_\parallel, Z)\left(\mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s\right)$ we cast Eq.(1.21) into the inhomogeneous $1D$ Helmotz equation:

$$\frac{d^2G_0(\mathbf{k}_\parallel, Z)}{dz^2} + k_z^2\ G_0(\mathbf{k}_\parallel, Z) = -\delta(Z) \tag{1.22}$$

This takes us to the mathematical definition of the free scalar scalar 1D Green's function as the solution of the inhomogeneous Helmotz equation [1, 2]:

$$G_0(\mathbf{k}_\parallel, Z) = \frac{\mathtt{i}}{2k_z}e^{ik_z|Z|} \tag{1.23}$$

which is actually a one dimensional running wave. The connection to the physical meaning of the 3D free-space scalar green function is attained through Sommerfeld's identity (also known as Weyl Identity [17–19]), which states that:

$$\frac{e^{igR}}{R} = \frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{\mathtt{i}}{2k_z} e^{ik_z|Z|} \tag{1.24}$$

This solution represents a spherical wave propagating away from the source $\mathbf{r}'$, in our case it is propagating from the origin since the source is centered at $\mathbf{r}' = 0$. The solution for the Green Tensor can be represented likewise, Fourier transforming its spectral form back to the spatial form:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \frac{\mathtt{i}}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{e^{ik_z|Z|}}{k_z}\, \left[\, \mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s \,\right]. \tag{1.25}$$

This integral representation of the space domain GT, is known as a Sommerfeld Integral. That means we have solved the differential Eq.(1.8). The boundary conditions in free space correspond to the Radiation Condition, which states that if we add an infinitesimal absorption to the wave $e^{\pm \mathtt{i} kZ}$ we must choose the sing in the exponential so that the wave decays at far from the source $Z = \pm\infty$. We make sure this is fulfilled if we choose $\Im\{k_z\} > 0$ and correspondingly represent the solution as:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \frac{\mathtt{i}}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{e^{ik_z Z}}{k_z}\, \left[\, \mathbf{k}_p^{+}\mathbf{k}_p^{+} + \mathbf{k}_s\mathbf{k}_s \,\right] \quad Z > 0$$

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \frac{\mathtt{i}}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{e^{-ik_z Z}}{k_z}\, \left[\, \mathbf{k}_p^{-}\mathbf{k}_p^{-} + \mathbf{k}_s\mathbf{k}_s \,\right] \quad Z < 0$$

Finally we can attain the explicit form of spatial representation of the free space Green Tensor, by simply using the correspondence between direct and reciprocal space : $\mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s \leftrightarrow \hat{\mathbf{1}} + \nabla\nabla$ and changing the order of the differential operators under the integral sign:

$$\frac{\mathtt{i}}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{e^{ik_z|Z|}}{k_z}\, \left[\, \mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm} + \mathbf{k}_s\mathbf{k}_s \,\right] = \tag{1.26}$$

$$= \{\hat{\mathbf{1}} + \nabla\nabla\,\} \left[\frac{\mathtt{i}}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_{\parallel}\, e^{i\mathbf{k}_{\parallel}\cdot\mathbf{R}_{\parallel}}\, \frac{e^{ik_z|Z|}}{k_z}\right] \quad . \tag{1.27}$$

In the last equation the object in the curly brackets is a tensorial operator while, the object in the square brackets is the scalar free-space green function $G_0 = e^{igR}/4\pi R$, see Eq.(1.24), so that finally the free-space GT can be written as:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \hat{\mathbf{1}}\, \frac{e^{igR}}{4\pi R} + \nabla\nabla\left(\frac{e^{igR}}{4\pi R}\right). \tag{1.28}$$

A graphic representation of the component G_{xx} of the tensor is provided in Fig.1.1. To get the explicit form for the Green tensor we need a brief reminder of tensor properties. First of all consider the spherical polar coordinates basis in direct space: ($\mathbf{u}_R$,$\mathbf{u}_\theta$, $\mathbf{u}_\phi$), which would be nothing more than the local frame of reference of an observer that moves on the surface of a sphere. Now remember that as we move on the surface of a sphere the normal to the surface $\mathbf{u}_R$ suffers an instantaneous change accordingly equal to $\mathbf{u}_\theta$ if we move longitude-wise and $\sin\theta\mathbf{u}_\phi$ if we do it latitude-wise. This is can be deduced mathematically by computing the following relations:

$$\frac{\partial \mathbf{u}_R}{\partial \theta} = \mathbf{u}_\theta \qquad \frac{\partial \mathbf{u}_R}{\partial \phi} = \sin\theta\, \mathbf{u}_\phi \tag{1.29}$$

Now the first tensor we are going to deal with, is the gradient of a vector $\mathbf{a}$ which in this basis reads:

$$\nabla \mathbf{a} = \mathbf{u}_R \frac{\partial \mathbf{a}}{\partial R} + \frac{\mathbf{u}_\theta}{R}\frac{\partial \mathbf{a}}{\partial \theta} + \mathbf{u}_\phi \frac{1}{R\sin\theta}\frac{\partial \mathbf{a}}{\partial \phi} \tag{1.30}$$

Hence:

$$\nabla \mathbf{u}_R = \frac{\mathbf{u}_\theta \mathbf{u}_\theta}{R} + \frac{\mathbf{u}_\phi \mathbf{u}_\phi}{R} \tag{1.31}$$

Now that we have derived this property we can get an explicit form for Green's tensor. The first step is generating the gradient of the scalar Green function:

$$\nabla G_0(R) = \left(ig - \frac{1}{R} \right) \frac{e^{igR}}{4\pi R}\, \mathbf{u}_R \tag{1.32}$$

and notice it has radial components only. However due to Eq.(1.31), the Green tensor will also have transversal components coming in turn from the gradient of the vector $\nabla G_0(R)$:

$$\nabla\nabla G_0(R) = \frac{e^{igR}}{4\pi}\left[\mathbf{u}_R\mathbf{u}_R \left(\frac{2}{R^3} - \frac{2\mathtt{i}g}{R^2} - \frac{g^2}{R} \right) + (\mathbf{u}_\phi\mathbf{u}_\phi + \mathbf{u}_\theta\mathbf{u}_\theta) \left(\frac{\mathtt{i}g}{R^2} - \frac{1}{R^3} \right) \right]$$

Notice how this part of the Green tensor consists of both near-field $1/R^3$ and far field $1/R$ components. Rearranging the $1/R$ components far-field component we get:

$$\begin{aligned} \frac{\nabla\nabla G_0(R)}{g^2} &= -\mathbf{u}_R\mathbf{u}_R \frac{e^{igR}}{4\pi R} + \\ &+ \frac{e^{igR}}{4\pi}\left[\mathbf{u}_R\mathbf{u}_R \left(\frac{2}{g^2R^3} - \frac{2\mathtt{i}}{gR^2} \right) + (\mathbf{u}_\phi\mathbf{u}_\phi + \mathbf{u}_\theta\mathbf{u}_\theta) \left(\frac{\mathtt{i}}{gR^2} - \frac{1}{g^2R^3} \right) \right] \end{aligned} \tag{1.33}$$

so that plugging this into Eq.(1.28) we end up with:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \frac{e^{igR}}{4\pi}\left[\mathbf{u}_R\mathbf{u}_R\left(\frac{-2\mathtt{i}}{gR^2}+\frac{2}{g^2R^3}\right)+(\mathbf{u}_\phi\mathbf{u}_\phi+\mathbf{u}_\theta\mathbf{u}_\theta)\left(\frac{1}{R}+\frac{\mathtt{i}}{gR^2}-\frac{1}{g^2R^3}\right)\right]$$

While this form of the Green tensor is equivalent to the others in the literature [8, 20], it has the benefit of making it easy to recognize that in the far-field, where the $1/R$ term dominates, the Green tensor is strictly a transversal tensor. Accordingly it generates transversal electric fields only, when expanded in the local frame of reference of a distant observer. Moreover its expression in the far-field is clear-cut in terms of the scalar Green function:

$$\hat{\mathbf{G}}_0(\mathbf{R}\to\infty) = G_0(R)(\mathbf{u}_\phi\mathbf{u}_\phi+\mathbf{u}_\theta\mathbf{u}_\theta) \tag{1.34}$$

As we shall see this general property may be useful to calculate the Poynting vector energy flux scattered by nano-defects.
A full account of tensor algebra (also referred to as dyadics), can be found in [2, 21].

1.3 The Green Tensor Divergence

As pointed out in Eq.(1.11) the Green's tensor contains a divergence at $R = 0$, which is now evident from Eq.(1.28) or Eq.(1.34). Indeed both the Green function and the Green Tensor are expected to have such divergences, since they are designed to reproduce the emission of a unit source $\delta(\mathbf{R})$ localized at $R = 0$. The divergent behavior at the origin, is also perceivable from the picture in Fig.1.1. Nonetheless some degree of ambiguity comes about with the Green Tensor when we carry out the integration of Eq.(1.5). There we have a volume integration about a singular static field that decays like $1/R^3$ and this is an improper integral. Handling this divergence has been the topic of extensive research[11, 15, 16, 21–24] over several decades, both in the Green Tensor Approach and in Diffraction theory[25]. In this Section and the next we present our own intuitive interpretation on what is the physical significance of the solution, which is justified rigorously in [15, 26]. Our derivation tackles the practical implementation of the formalism. Specifically, the problem of how to apply Eq.(1.28) in the middle of a current-carrying beam, raised in[27].
The Green tensor is: $G_0(\mathbf{R}) = \hat{\mathbf{1}} \, \frac{e^{igR}}{4\pi R} \, + \nabla\nabla \left(\frac{e^{igR}}{4\pi R} \right)$. Now, if we isolate an infinitesimal volume V_δ , which is taken arbitrarily small around the origin:

$$\mathbf{E}_s(\mathbf{r}) = \int_{V-V_\delta} d\mathbf{r}' \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{\Theta}(\mathbf{r}') + \left[\lim_{V_\delta \to 0} \int_{V_\delta} d\mathbf{r}' \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \right] \cdot \mathbf{\Theta}(\mathbf{r}) \quad (1.35)$$

hence :

$$\lim_{V_\delta \to 0} \int_{V_\delta} d\mathbf{r}' \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') = \frac{1}{g^2} \lim_{V_\delta \to 0} \int_{V_\delta} d\mathbf{R}' \nabla(\nabla G_0(R)) \quad (1.36)$$

This term can be turned into a proper integral by using the dyadic divergence theorem [21]:

$$\int_V d\mathbf{R} \, \nabla \mathbf{a} \, = \oint_S dS \, \mathbf{n} \, \mathbf{a} \quad (1.37)$$

Using Eq.(1.32) and Eq.(1.37) we have:

$$\int_V d\mathbf{R}' \, \nabla(\nabla G_0(R)) = \oint_S dS \, \mathbf{n} \, \nabla G_0(R) = -\frac{1}{4\pi} \oint dS \, \frac{\mathbf{n} \, \mathbf{u}_R}{R^2} \quad (1.38)$$

where we have dropped the first term of Eq.(1.32) because when taking the limit $\oint_S dS \; \mathbf{n} \mathbf{u}_R \mathtt{i} g/R$ over a vanishing surface, this term also vanishes. In fact the integration over the surface goes like R^2 while the integrand decays like R^{-1}. On the contrary the second term of Eq.(1.32) has produced $\frac{1}{4\pi} \oint dS \; \mathbf{n} \, \mathbf{u}_R / R^2$, which is independent of the size of the exclusion volume and only depends on its shape and orientation, as in fact it can be interpreted

mathematically as a normalized solid angle tensor[15]. Now we are in a position to redefine the improper integral as:

$$\mathbf{E}_s(\mathbf{r}) = g^2 \int_{V-V_\delta} d\mathbf{r}' \hat{\mathbf{G}}(\mathbf{r},\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}') - \left[\frac{1}{4\pi} \oint dS \, \frac{\mathbf{n}\,\mathbf{u}_R}{R^2}\right] \cdot \mathbf{P}(\mathbf{r}) \tag{1.39}$$

Hence we shall designate:

$$\hat{\mathbf{L}}(\mathbf{r}) = \left[\frac{1}{4\pi} \oint dS \, \frac{\mathbf{n}\,\mathbf{u}_R}{R^2}\right] \tag{1.40}$$

In conclusion the term $\hat{\mathbf{L}}(\mathbf{r})$ takes care of the singularity of the Green Tensor mathematically, and is known as the depolarization tensor. Next we are going to focus on the depolarization tensor and expand on its physical origin.

1.4 Depolarizing field and Depolarization dyadic

The Green Tensor Approach is a compact formulation of electromagnetism whose hallmark is condensing field's properties and boundary conditions useful in electrodynamic problems. Nevertheless it also provides an alternative summary of laborious electrostatic problems. I will now revise the relation between depolarization and polarization and give our own interpretation of the physical significance of the last term of Eq.(1.39).

When a polarizable object is immersed in an electrostatic field $\mathbf{E}_b(\mathbf{r})$ the field inside the object can be expressed as a superposition of the applied field and the incremental field produced by the polarization induced on the object. The latter can be written as a linear combination of the components of the polarization $\mathbf{P}(\mathbf{r})$, and the corresponding coefficients are named 'depolarization factors'. Usually the electrostatic problem is too complicated and it is only solved for symmetrical defects. For example in [28] it is solved for a sphere where, due to its symmetry, you have equal depolarization factor in all directions.

Now let us go back to the Green tensor approach. Notice that in Eq.(1.39) if we consider a volume $V = V_\delta$, the particular solution $\mathbf{E}_s(\mathbf{r})$ we have found reduces to:

$$\mathbf{E}_s(\mathbf{r}) = -\left[\frac{1}{4\pi} \oint dS \, \frac{\mathbf{n}\,\mathbf{u}_R}{R^2}\right] \cdot \mathbf{P}(\mathbf{r}) \tag{1.41}$$

This corresponds to the long-wavelength limit of Maxwell's equations in which the incident wavelength is very much larger than the volume of the scatterer. The scatterer will see the field as a static field. As the volume is infinitesimal, light takes no time in crossing it, so retardation is negligible. Therefore the full solution to Eq.(1.3) is attained by adding to it the homogenous solution $\mathbf{E}_b(\mathbf{r})$ as:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_b(\mathbf{r}) + \mathbf{E}_s(\mathbf{r}) \tag{1.42}$$

or in terms of the depolarization tensor:

$$\mathbf{E}(\mathbf{r}) - \mathbf{E}_b(\mathbf{r}) = -\hat{\mathbf{L}}(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}). \tag{1.43}$$

The term $\mathbf{E}(\mathbf{r}) - \mathbf{E}_b(\mathbf{r}) = \mathbf{E}_s(\mathbf{r})$ is known as the depolarizing field, it can also be thought of as the perturbation to the background field arising from the polarization charges which position themselves on the surface of the object V. In a static field this charge distribution has no relation to the size of the object and that is why the integral representing $\hat{\mathbf{L}}(\mathbf{r})$ is independent of the size of V_δ. We see that $\hat{\mathbf{L}}(\mathbf{r})$ represents physically the depolarization tensor since each of its components is a depolarization factor. $\mathbf{L}$ depends univocally on the shape and orientation of V.
Mathematically $\hat{\mathbf{L}}$ has various properties proved in [15]. Particularly it has a unitary trace, namely:

$$0 < L_{ii} < 1 \quad ; \quad \sum_i L_{ii} = 1. \tag{1.44}$$

The depolarizing field is the dot product of an equivalent dipole placed at $\mathbf{r}$ with moment $\mathbf{P} = (\varepsilon - 1)\mathbf{E}$ and the depolarizing dyadic. This substantially tells us where an equivalent dipole stands with respect to the field lines. Once we have fixed the material permittivity, the net electromagnetic properties are determined by the shape and orientation of the surface of the object. A one

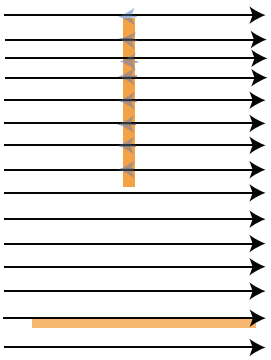

Figure 1.2: A needle-shaped defect in a static electric field. When the needle is vertical to the field lines a depolarizing field is established to counter the applied field. When the needle is parallel to the field lines there is no depolarizing field.

dimensional inhomogeneity can help visualize how the value of $\hat{\mathbf{L}}(\mathbf{r})$ relates to the depolarizing field for an object of general shape and orientation, since then $\hat{\mathbf{L}}(\mathbf{r})$ can be deduced from the boundary conditions of the fields alone. Let us consider two possible orientations of the needle-shaped polarizable object. With reference to Fig.1.2:
i) When the needle has its long side perpendicular to the field we have: $\mathbf{D} =$

const or $E_b(\mathbf{r}) = \varepsilon E(\mathbf{r}) = (\varepsilon - 1)E(\mathbf{r}) + E(\mathbf{r}) = E(\mathbf{r}) + P(\mathbf{r})$. This means $L = 1$. In accordance with the unity trace of $\mathbf{L}$ we have the maximum variation of the electric field E from the reference field. Accordingly $E(\mathbf{r}) = E_b(\mathbf{r})/\varepsilon$; the applied field inside V cannot be reduced more than the permittivity ε of V, because this corresponds to the maximum storage of the field energy.
ii) when the needle has its long side parallel to the applied field lines, the continuity of $E_{\|}$ means the field inside the needle is equal to the applied field. Accordingly we have $L = 0$ which represents no variation of the reference field inside the object.
An object of arbitrary shape and orientation has at least two dimensions with $L_{ii} \neq 0$, and thus in general none of the components can actually be equal to 1. However we can interpret (1.43) by saying that, the variation of the internal field from the applied field in any direction will always be something in-between the two cases exemplified by the two geometries of the needle. The depolarizing factor in a given direction, given by the relative components of $\hat{\mathbf{L}}(\mathbf{r})$ must be a fraction contained between 0 and 1, and thus may be intuitively understood as a projection of the curvature of the surface of the object on the shape of a flat needle oriented in that direction.

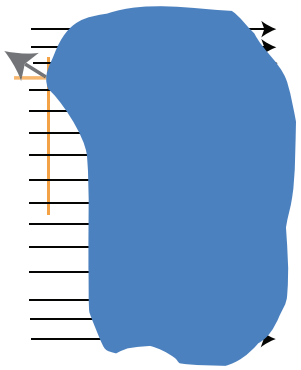

Figure 1.3: The normal at a point on the surface of a scatterer of arbitrary shape can be resolved into orthogonal components parallel and perpendicular to the applied field. Each of these components is a fraction of the contribution that two orthogonal needle type defects would give to the depolarization of the object.

1.5 The Electromagnetic Lippmann-Schwinger equation in a homogenous background

We now have all the pieces of the solution for the scattering of light by an arbitrary scatterer of volume V and dielectric constant ε in free-space. In the last section we have solved the problem for an infinitesimally small object. If we go back to the general case of an arbitrary-sized volume V, we are in a position of writing the full solution to Eq.(1.3) as $\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) + \mathbf{E}_s(\mathbf{r})$ or:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) - \Delta\varepsilon\,\mathbf{L}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}) + g^2\,\Delta\varepsilon \int_{V-V_\delta} d\mathbf{r}'\hat{\mathbf{G}}_0(\mathbf{r},\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}') \quad (1.45)$$

We have derived it for free space background. However it is easy to generalize it to an inhomogeneities of volume V and dielectric constant ε in a general homogeneous background medium with dielectric constant ε_b .
In this case the dielectric contrast between the homogenous background with dielectric and the inhomogeneity is: $\Delta\varepsilon = \varepsilon - \varepsilon_b$. The whole derivation is the same if we substitute g for $g\sqrt{\varepsilon_b}$ in the Green tensor:

$$\hat{\mathbf{G}}_b(\mathbf{r}-\mathbf{r}') = \left(\hat{\mathbf{1}} + \frac{\nabla\nabla}{\varepsilon_b g^2}\right) \frac{e^{\mathrm{i}\sqrt{\varepsilon_b}\,gR}}{4\pi R} \quad (1.46)$$

so that in Eq.(1.39) we would have:

$$\mathbf{E}_s(\mathbf{r}) = g^2 \int_{V-V_\delta} d\mathbf{r}'\hat{\mathbf{G}}_b(\mathbf{r}-\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}') - \left[\frac{1}{4\pi}\oint dS\,\frac{\mathbf{n}\,\mathbf{u}_R}{R^2}\right] \cdot \frac{\mathbf{P}(\mathbf{r})}{\varepsilon_b}. \quad (1.47)$$

to yield the Lippmann-Schwinger Equation:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_b(\mathbf{r}) - \Delta\varepsilon\frac{\mathbf{L}(\mathbf{r})}{\varepsilon_b} \cdot \mathbf{E}(\mathbf{r}) + g^2\,\Delta\varepsilon \int_{V-V_\delta} d\mathbf{r}'\hat{\mathbf{G}}_b(\mathbf{r}-\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}'). \quad (1.48)$$

1.6 Discretization

In numerical computation Eq.(1.48) can be be solved by creating a computer model of the scatterer wherein we discretize the volume V of the inhomogeneity into an array of N meshes of volume V_i. In turn, this requires extracting one further correcting static term. This is because when we solve the integral in Eq.(1.48) by discretizing it, we are approximating an infinitesimal volume with a small finite volume: $V_i \sim V_\delta$. Hence we have to account for the difference between the volume over the mesh V_i and V_δ through the correction:

$$\hat{\mathbf{M}}(\mathbf{r}_i) = \lim_{V_\delta \to 0} \int_{V_i - V_\delta} d\mathbf{r}'\hat{\mathbf{G}}(\mathbf{r}_i,\mathbf{r}')d\mathbf{r}' \quad (1.49)$$

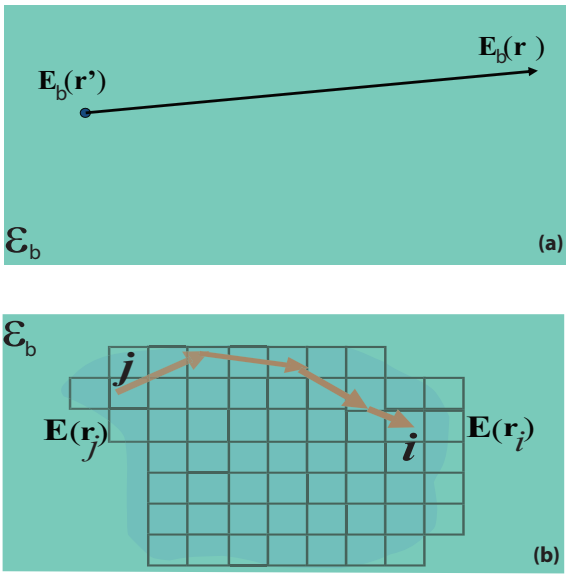

Figure 1.4: A schematic representation of a homogeneous (infinite) back-ground where light is not diffracted (Panel (a)) and of the the inhomogeneity in a homogeneous background (Panel (b)). The inhomogeneity is discretized. The electric field of any mesh j propagates directly to any another mesh i. However light can propagate on any path between the two points.

where $V_i = V/N^3$. If we do that, we yield the discretized Lippmann-Schwinger:

$$\mathbf{E}(\mathbf{r}_i) = \mathbf{E}_b(\mathbf{r}_i) \; + \; g^2 \, \Delta\varepsilon \sum_{j \neq i} \hat{\mathbf{G}}_b(\mathbf{r}_i - \mathbf{r}_j) \cdot \mathbf{E}(\mathbf{r}_j) V_j$$
$$+ g^2 \, \Delta\varepsilon \hat{\mathbf{M}}(\mathbf{r}_i) \cdot \mathbf{E}(\mathbf{r}_i) - \Delta\varepsilon \frac{\hat{\mathbf{L}}(\mathbf{r}_i)}{\varepsilon_b} \cdot \mathbf{E}(\mathbf{r}_i) \tag{1.50}$$

where $\mathbf{E}(\mathbf{r}_i)$ is the field at the center of one mesh and $\hat{\mathbf{L}}(\mathbf{r}_i)$ is the depolarization tensor calculated at the center of the mesh. Its value, from [15], is: $\hat{\mathbf{L}}(\mathbf{r}_i) = \hat{\mathbf{1}}\, 1/3$.
While the term $\hat{\mathbf{L}}(\mathbf{r}_i)$ comes from the part of the Green Tensor Eq.(1.34) that goes like $1/R^3$, the contribution $\hat{\mathbf{M}}(\mathbf{r}_i)$ comes from the term of the Green Tensor that goes like $1/R$. This is because the terms that go like $1/R^2$ and $1/R^3$ cancel out, when we integrate the Green Tensor over the angles in spherical polar coordinates. Take in fact the Green Tensor:

$$\hat{\mathbf{G}}_0(\mathbf{R}) = \frac{e^{igR}}{4\pi} \left[\mathbf{u}_R \mathbf{u}_R \left(\frac{2}{g^2 R^3} - \frac{2\mathtt{i}}{gR^2} \right) + (\mathbf{u}_\phi \mathbf{u}_\phi + \mathbf{u}_\theta \mathbf{u}_\theta) \left(\frac{1}{R} + \frac{\mathtt{i}}{gR^2} - \frac{1}{g^2 R^3} \right) \right] \tag{1.51}$$

If we use the identities derived in the Appendix:

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \, \mathbf{u}_R \mathbf{u}_R = \frac{4\pi}{3} \hat{\mathbf{1}} \tag{1.52}$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \, (\mathbf{u}_\theta \mathbf{u}_\theta + \mathbf{u}_\phi \mathbf{u}_\phi) = \frac{8\pi}{3} \hat{\mathbf{1}} \tag{1.53}$$

we get that, after integrating $2/(g^2R^3)\mathbf{u}_R\mathbf{u}_R$ over both ϕ and θ, this cancels the same integral integral of $-1/(g^2R^3)(\mathbf{u}_\phi\mathbf{u}_\phi + \mathbf{u}_\theta\mathbf{u}_\theta)$. At the same time, the integral over both ϕ and θ of the term $-2\mathtt{i}/(gR^2)\mathbf{u}_R\mathbf{u}_R + \mathtt{i}/(gR^2)(\mathbf{u}_\phi\mathbf{u}_\phi + \mathbf{u}_\theta\mathbf{u}_\theta)$ is zero. Finally we are left with: $\frac{e^{igR}}{4\pi R}(\mathbf{u}_\phi\mathbf{u}_\phi + \mathbf{u}_\theta\mathbf{u}_\theta)$. This is the only term that does not cancel and:

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \, \hat{\mathbf{G}}_0(\mathbf{R}) = \frac{2}{3}\frac{e^{igR}}{R}\hat{\mathbf{1}}. \tag{1.54}$$

Lastly we set: $\mathbf{r}_i = 0$ and $R = r'$ and consequently we can complete the volume integration by integrating over R to get:

$$\hat{\mathbf{M}} = \frac{2}{3}\,\hat{\mathbf{1}} \int_0^R dR'\, R'\, e^{igR'} = \frac{2}{3g^2}\,\hat{\mathbf{1}}\,\left[(1 - \mathtt{i}gR)e^{\mathtt{i}gR} - 1\right] \approx \hat{\mathbf{1}}\,\frac{R^2}{3} \tag{1.55}$$

This is our own derivation. An alternative and old derivation of this static term of the Green Tensor can be found in [29]. More recent reviews like [8], prescribe the use Eq.(1.55) in numerical evaluation, by setting $R = R_i$ where that latter is defined by: $V_i = 4\pi R_i^3/3$.
Therefore in the discrete version the Maxwell equations the volume V is represented by an array of N points. At each point $\mathbf{r}_i$ we impose Eq.(1.50) and generate a set of N-coupled equations. This set can be solved by the standard Conjugate Gradients algorithm. A pedagogical implementation of this method can be found in [30].

1.7 Lipmann-Schwinger Equation in a the presentence of an interface

For a semi-space background we need to add to the radiation condition which represent the boundary condition of free space, the boundary conditions of the flat surface. This can be achieved by defining the total Green Tensor:

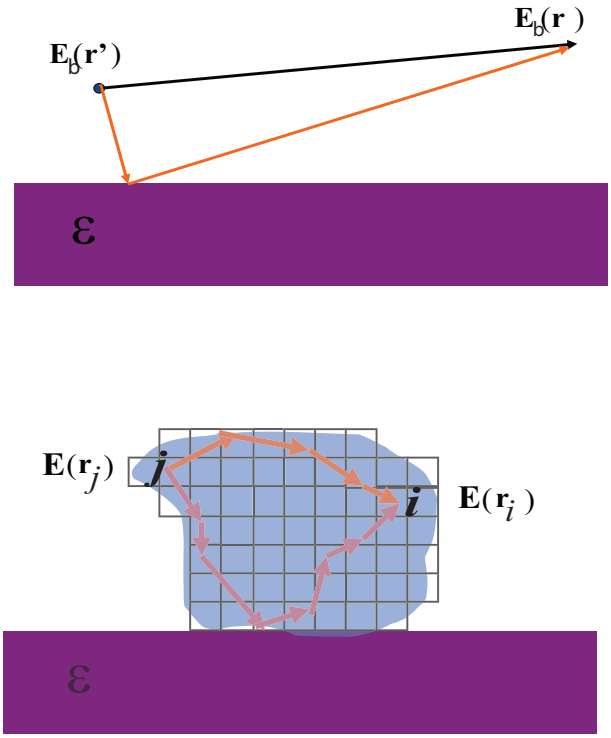

Figure 1.5: (Panel (a)) A schematic representation of a semi-infinite back-ground. Light propagates directly between two points, or indirectly after being reflected by the interface. (Panel (b)) Light is diffracted by a discretized inhomogeneity The electric field of any mesh j propagates to any another mesh i on any path connecting the two points either directly or indirectly (after reflection at the interface).

$$\hat{\mathbf{G}}(\mathbf{r},\mathbf{r}') = \hat{\mathbf{G}}_b(\mathbf{r}-\mathbf{r}') + \hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}') \tag{1.56}$$

where by definition $\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')$ is called the *indirect Green Tensor* and is a complementary solution of :

$$\nabla \times \nabla \times \hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}') - g^2\, \hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}') = 0 \tag{1.57}$$

such that $\hat{\mathbf{G}}(\mathbf{r},\mathbf{r}')$ satisfies the new background boundary conditions at the surface. Such method is the method of scattering superposition and is developed in the next section.

Once $\hat{\mathbf{G}}(\mathbf{r},\mathbf{r}')$ is obtained in the semispace background the field inside the scatterer can be attained by solving:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_b(\mathbf{r}) - \Delta\varepsilon\frac{\mathbf{L}(\mathbf{r})}{\varepsilon_b}\cdot\mathbf{E}(\mathbf{r}) + g^2\,\Delta\varepsilon\int_{V-V_\delta} d\mathbf{r}'\hat{\mathbf{G}}(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}(\mathbf{r}') \tag{1.58}$$

which can be descretized to:

$$\begin{aligned}\mathbf{E}(\mathbf{r}_i) = \mathbf{E}_b(\mathbf{r}_i)\ &+\ g^2\sum_{j\neq i}\hat{\mathbf{G}}_b(\mathbf{r}_i,\mathbf{r}_j)\cdot\Delta\varepsilon\,\mathbf{E}(\mathbf{r}_j)V_j + \\ &+g^2\,\hat{\mathbf{M}}\cdot\Delta\varepsilon\,\mathbf{E}(\mathbf{r}_i) - \frac{\hat{\mathbf{L}}}{\varepsilon_b}\cdot\Delta\varepsilon\,\mathbf{E}(\mathbf{r}_i) + \\ &+g^2\sum_j\hat{\mathbf{G}}_I(\mathbf{r}_i,\mathbf{r}_j)\cdot\Delta\varepsilon\,\mathbf{E}(\mathbf{r}_j)V_j\end{aligned} \tag{1.59}$$

1.8 The Green Tensor for a semi-space: The Method of scattering superposition

The following derivation is our own, except for parts where references are cited. In order to find the Green tensor additional term $\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')$ which takes care of the interaction with an interface, we need to reproduce the path that led to the derivation of the free-space GT through its spectral representation. We started with the space equation:

$$\nabla\times\nabla\times\hat{\mathbf{G}}_0(\mathbf{r}-\mathbf{r}')-g^2\,\hat{\mathbf{G}}_0(\mathbf{r}-\mathbf{r}')=\delta(\mathbf{r}-\mathbf{r}') \tag{1.60}$$

and transformed it into:

$$\frac{d^2\hat{\mathbf{G}}_0(\mathbf{k}_\parallel,Z)}{dz^2}+k_z^2\,\hat{\mathbf{G}}_0(\mathbf{k}_\parallel,Z)=-\left(\mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm}+\mathbf{k}_s\mathbf{k}_s\right)\delta(Z) \tag{1.61}$$

Then we assumed there exists a scalar function $\hat{\mathbf{G}}_0(\mathbf{k}_\parallel,Z)$ such that: $\hat{\mathbf{G}}_0(\mathbf{k}_\parallel,Z)=G_0(\mathbf{k}_\parallel,Z)\left(\mathbf{k}_p^{\pm}\mathbf{k}_p^{\pm}+\mathbf{k}_s\mathbf{k}_s\right)$ to cast Eq.(1.61) into the inhomogeneous $1D$ Helmotz equation:

$$\frac{d^2G_0(\mathbf{k}_\parallel,Z)}{dz^2}+k_z^2\,G_0(\mathbf{k}_\parallel,Z)=-\delta(Z) \tag{1.62}$$

Finding $\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')$ will be even easier with this technique. Let us assume the metal plane interface separates a dielectric background ε_b from a metal background ε_m and set this on the xy-plane at $z=0$. We also assume that the semi-infinite region $z>0$ is filled with a dielectric and $z<0$ is filled with the metal semi-space. Since all dielectric constants are uniform and isotropic, the space symmetry is broken in the z direction but is preserved in the xy-direction. Hence $\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')=\hat{\mathbf{G}}_I(\mathbf{r}_\parallel-\mathbf{r}'_\parallel,z,z')$ so that :

$$\nabla\times\nabla\times\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')-g^2\,\hat{\mathbf{G}}_I(\mathbf{r},\mathbf{r}')=0 \tag{1.63}$$

may Fourier transformed into:

$$\frac{d^2\hat{\mathbf{G}}_I(\mathbf{k}_\parallel,z,z')}{dz^2}+k_z^2\,\hat{\mathbf{G}}_I(\mathbf{k}_\parallel,z,z')=0 \tag{1.64}$$

From the last equation we attain the contribution of the metal plane interface. This can be split into the p-polarized waves and s-waves, which we have already introduced. Yet the two modes have different reflection coefficients at the interface. In fact their vectorial structure is different due to their polarization with respect to the interface. For a p-wave (transverse magnetic mode) the sign of its components depends on the direction of propagation. For an s-wave we do not make such distinction, because its components are independent of

the direction of propagation, being these transverse electric modes. The actual vectors $\mathbf{k}_p^{\pm}$ and $\mathbf{k}_s$ are the vectorial part of the a normalized plane wave mode of either polarization.
Consider now the concrete case of a point in the dielectric semi-space. $\hat{\mathbf{G}}_I(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z, z')$ is an operator whose function is to propagate a wave that travels from

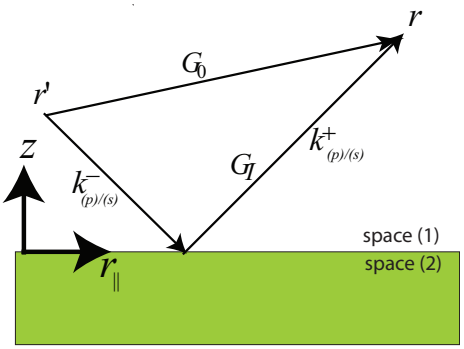

Figure 1.6: A schematic representation of $\hat{\mathbf{G}}_0$ and $\hat{\mathbf{G}}_I$ for a vacuum-metal background. $\hat{\mathbf{G}}_0$ propagates the light emitted by a source at $\mathbf{r}'$ directly to $\mathbf{r}$. $\hat{\mathbf{G}}_I$ propagates the light emitted by a source at $\mathbf{r}'$ indirectly to $\mathbf{r}$, by reflection at the interface. The direction of propagation of the modes $\mathbf{k}^{\pm}_{(p)/(s)}$ is also illustrated.

its source at a point $z' > 0$, to the interface where it is reflected onto the point $z > 0$. Hence, in this case the Green Tensor should operate on a wave propagating towards the interface, that is in the negative direction, and turn it into a wave propagating from the surface outwards, i.e. into the positive direction. Particularly this has to be imposed in the vectorial structure of p-waves, and the relevant part of the tensor $\hat{\mathbf{G}}_I$ must go like: $\mathbf{k}_p^- \mathbf{k}_p^+$. On the whole, let us assume $\hat{\mathbf{G}}_I(\mathbf{k}_{\parallel}, z, z') = {G_I}^{(p)} \mathbf{k}_p^- \mathbf{k}_p^+ + {G_I}^{(s)} \mathbf{k}_s \mathbf{k}_s$ and we get:

$$\frac{d^2 G_I^{(p/s)}(\mathbf{k}_{\parallel}, z, z')}{dz^2} + k_z^2 \, {G_I}^{(p/s)}(\mathbf{k}_{\parallel}, z, z) = 0. \tag{1.65}$$

As explained in [2] the solutions are:

$${G_I}^{(p)}(\mathbf{k}_{\parallel}, z, z) = \frac{\mathtt{i}}{2k_z} r_p^{(1,1)} \; e^{ik_z(z+z')} \tag{1.66}$$

$${G_I}^{(s)}(\mathbf{k}_{\parallel}, z, z) = \frac{\mathtt{i}}{2k_z} r_s^{(1,1)} \; e^{ik_z(z+z')} \tag{1.67}$$

and consequently, if we define $z + z' = Z_+$, we can write:

$$\hat{\mathbf{G}}_I(\mathbf{R}_{\parallel}, Z_+) = \frac{i}{8\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{k}_{\parallel} \, e^{i\mathbf{k}_{\parallel} \cdot \mathbf{R}_{\parallel}} \, \frac{i}{2k_z} e^{ik_z Z_+} \left[\, r_p^{(1,1)} \mathbf{k}_p^- \mathbf{k}_p^+ + r_s^{(1,1)} \mathbf{k}_s \mathbf{k}_s \, \right]. \tag{1.68}$$

Notice $r_p^{(1,1)}$ and $r_s^{(1,1)}$ are the Fresnel reflection coefficients which we are going to deduce below by matching the electric field boundary conditions at the plane interface XY set at $z' = 0$. Before we do that we need to define plane waves on the metal side of the interface. As said p-waves are refracted at the plane interface and thus need to be defined in the background medium ε_b as:

$$\mathbf{k}_p^{b\pm} = \frac{k_z^b}{g k_\parallel}\left(k_x \mathbf{u}_x + k_y \mathbf{u}_y \mp \frac{k_\parallel^2}{k_z^b}\mathbf{u}_z\right) \tag{1.69}$$

where

$$k_z^b = \sqrt{\varepsilon_b g^2 - k_\parallel^2}. \tag{1.70}$$

On the other hand s-waves are continuous though the plane interface.
In this treatment we are only going to consider two media.
i) Free-space: $\varepsilon_b = 1$ and correspondingly the p-modes will be designated simply: $\mathbf{k}_p$ and k_z^b simply k_z.
ii) Metal $\varepsilon_b = \varepsilon$ and we shall designate $\mathbf{k}_p^m$ and k_z^m. We can now proceed to define the Fresnel reflection and transmission coefficients. These are attained by writing the net fields generated at either sides of the interface from an incident wave in the two polarizations. For p-waves:

$$\mathbf{E}_p^{(1,1)}(z') = e^{-ik_z z'}\mathbf{k}_p^- + e^{ik_z z'} r_p^{(1,1)} \mathbf{k}_p^+ \tag{1.71}$$

$$\mathbf{E}_p^{(1,2)}(z') = e^{-ik_z^m z'}\, t_p^{(1,2)} \mathbf{k}_p^{m-} \tag{1.72}$$

or an incident s-wave:

$$\mathbf{E}_s^{(1,1)}(z') = e^{-ik_z z'}\mathbf{k}_s + e^{ik_z z'} r_s^{(1,1)} \mathbf{k}_s \tag{1.73}$$

$$\mathbf{E}_s^{(1,2)}(z') = e^{-ik_z^m z'}\, t_s^{(1,2)} \mathbf{k}_s \tag{1.74}$$

to yield:

$$r_p^{(1,1)} = \frac{k_z^m - \varepsilon\, k_z}{k_z^m + \varepsilon\, k_z} \tag{1.75}$$

$$t_p^{(1,2)} = \frac{2k_z\sqrt{\varepsilon}}{k_z^m + \varepsilon\, k_z}. \tag{1.76}$$

1.9 2D Systems

A large part of this thesis will deal with 2D systems, that is surface defects that are infinite in one of the dimensions parallel to the surface. This shall always be designated as the y-direction. The relative formalism is just an analog to the 3D formalism and most of the derivations are redundant. In the next section we shall quickly show how to convert the 3D formalism just developed, in the relevant 2D framework adopted.

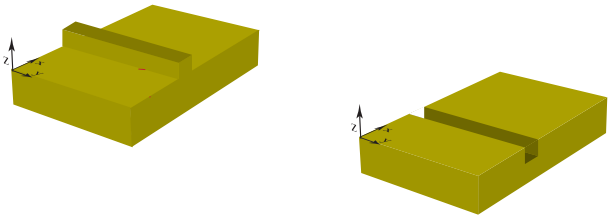

Figure 1.7: A schematic representation of 2D surface defects. These are defects that are infinite in the y-direction and can be either protrusions or indentations.

This problem has cylindrical symmetry and consequently we adopt cylindrical coordinates: $\boldsymbol{\rho} = x\mathbf{u}_x + z\mathbf{u}_z$. In this case we make a simplifying assumption on the background field and that is:

$$\mathbf{E}_b(\mathbf{r}) = \mathbf{E}_b(\boldsymbol{\rho})\, e^{ik_y^0 y} \tag{1.77}$$

To find the electric field in this case we need to go back to Eq.(1.6) and carry out the integration over the y-axis. We can find a particular solution of the inhomogeneous equation Fourier transforming the 3D solutions. Consider:

$$\mathbf{E}_s(\mathbf{r}) = g^2 \int_V d\mathbf{r}'\, \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{P}(\mathbf{r}') \tag{1.78}$$

and define :

$$\mathbf{E}_s(\boldsymbol{\rho}) = \int_{-\infty}^{\infty} dy\, e^{ik_y^0 y}\, \mathbf{E}_s(\mathbf{r}) \tag{1.79}$$

Hence given the cross-sectional area of the defect, A, the solution will be :

$$\mathbf{E}_s(\boldsymbol{\rho}) = g^2 \int_A \Delta\varepsilon\, d\mathbf{r}'\, \hat{\mathbf{G}}_{2D}(\boldsymbol{\rho}, \boldsymbol{\rho}') \cdot \mathbf{E}(\boldsymbol{\rho}') \tag{1.80}$$

with:

$$\hat{\mathbf{G}}_{2D}(\boldsymbol{\rho}, \boldsymbol{\rho}') = \int_{-\infty}^{\infty} dy\, e^{ik_{y0} y}\, \hat{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \tag{1.81}$$

Where we used the translational symmetry that the Green Tensor has on the plane. This is explicit in its spectral representation that, as explained, takes the form:

$$\hat{\mathbf{G}}(\mathbf{r},\mathbf{r}') = \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \; e^{ik_x(x-x')} e^{ik_y(y-y')} \, \hat{\mathbf{G}}(k_x, k_z, z, z') \qquad (1.82)$$

in a semi-space background.
To sum up, as a result the symmetry of the scattering system in the y direction the momentum in this direction is conserved and the component of the incident wavevector k_y^0 is cyclic. Consequently all fields depend on the y direction by a phase factor $e^{ik_y^0 y}$ which we will omit, but it is considered implicit in the equations. Also we shall drop the superscript k_y^0 when talking about two dimensional systems and refer to it simply as k_y, but it is meant it is the value of the y component of the incident wavevector.

1.10 Homogeneous medium Dyadic Green's Function

In this section we are going to deduce the 2D Green Tensor for an arbitrary background ε_b. We have seen in Sec.1.5 that the free space Green Tensor is converted to the homogenous medium Green Tensor by substituting g for $\sqrt{\varepsilon_b}\ g$. This time to we are going to condense our notation so that it holds in any homogeneous background, by defining the Green Tensor in terms of $\mathsf{g} = \sqrt{\varepsilon_b}\ g$. Once we have derived the Green tensor we will be able in the next sections to attain the relative depolarization dyadic $\mathbf{L}_{2D}$ and finally the correction term M_{2D} we need when discretizing the fields.
Applying Eq.(1.81) to the spectral representation of the GT we have:

$$\hat{\mathbf{G}}_b(X_-, Z_-) = \frac{\mathsf{i}}{4\pi}\int_{-\infty}^{\infty} dk_x\, e^{ik_x X_-}\, \frac{e^{ik_z Z_-}}{k_z}\ \left[\mathbf{k}_p^{b\pm}\mathbf{k}_p^{b\pm} + \mathbf{k}_s\mathbf{k}_s\right] \tag{1.83}$$

where $X_- = x - x'$ and $Z_- = z - z'$. Notice that the whole integrand is just a function of k_x . Now if we write explicit the dyadic in the integrand we have:

$$\mathbf{k}_p^{b\pm}\mathbf{k}_p^{b\pm} + \mathbf{k}_s\mathbf{k}_s = \begin{pmatrix} 1 - \frac{k_x^2}{\mathsf{g}^2} & \frac{-k_x k_y}{\mathsf{g}^2} & \frac{-k_x k_z^b}{\mathsf{g}^2} \\ \frac{-k_x k_y}{\mathsf{g}^2} & 1 - \frac{{k_y}^2}{\mathsf{g}^2} & \frac{-k_y k_z^b}{\mathsf{g}2} \\ \frac{-k_x k_z^b}{\mathsf{g}^2} & \frac{-k_y k_z^b}{\mathsf{g}^2} & 1 - \frac{k_z^{b2}}{\mathsf{g}^2} \end{pmatrix} \tag{1.84}$$

and define the following tensorial operator:

$$\hat{\Lambda} = \begin{pmatrix} 1 + \frac{\partial_x^2}{\mathsf{g}^2} & \frac{ik_y\partial_x}{\mathsf{g}^2} & \frac{\partial_x\partial_z}{\mathsf{g}^2} \\ \frac{ik_y\partial_x}{\mathsf{g}^2} & 1 - \frac{k_y^2}{\mathsf{g}^2} & \frac{ik_y\partial_x}{\mathsf{g}^2} \\ \frac{\partial_x\partial_z}{\mathsf{g}^2} & \frac{ik_y\partial_z}{\mathsf{g}^2} & 1 + \frac{\partial_z^2}{\mathsf{g}^2} \end{pmatrix} \tag{1.85}$$

If we then designate $\mathsf{R} = X_- + Z_-,\ \kappa = \sqrt{\mathsf{g}^2 - {k_y}^2}$, $\gamma = \kappa/\mathsf{g}$ and use the 2D version of Sommerfeld's identity:

$$\pi H_0(\kappa\mathsf{R}) = \int_{-\infty}^{\infty} \frac{dq}{\sqrt{1-q^2}} e^{i\kappa\sqrt{1-q^2}Z_-}\, e^{i\kappa q X_-} \tag{1.86}$$

we end up with:

$$\hat{\mathbf{G}}_b(\mathbf{R}\ ) = \frac{i}{4}\hat{\Lambda}\, H_0(\kappa\mathsf{R}). \tag{1.87}$$

From here the components of $\hat{\mathbf{G}}_b(\mathbf{R}\ )$ can be found explicitly as:

$$G_{bxx}(\mathbf{R}\ ) = \frac{i}{4}H_0(\kappa\mathsf{R})\left(1-\frac{\gamma^2}{2}\right)+\frac{i}{4}\gamma^2 H_2(\kappa\mathsf{R})\left(\left(\frac{X_-}{\mathsf{R}}\right)^2-\frac{1}{2}\right) \quad (1.88)$$

$$G_{byy}(\mathbf{R}\ ) = \frac{i}{4}\left(1-\frac{k_y^2}{\mathsf{g}^2}\right)H_0(\kappa\mathsf{R}) \quad (1.89)$$

$$G_{bzz}(\mathbf{R}\ ) = \frac{i}{4}H_0(\kappa\mathsf{R})\left(1-\frac{\gamma^2}{2}\right)+\frac{i}{4}\gamma^2 H_2(\kappa\mathsf{R})\left(\left(\frac{Z_-}{\mathsf{R}}\right)^2-\frac{1}{2}\right) \quad (1.90)$$

$$G_{bxy}(\mathbf{R}\ ) = \frac{\gamma}{4}\frac{k_y}{\mathsf{g}}H_1(\kappa\mathsf{R})\frac{X_-}{\mathsf{R}} \quad (1.91)$$

$$G_{byz}(\mathbf{R}\ ) = \frac{\gamma}{4}\frac{k_y}{\mathsf{g}}H_1(\kappa\mathsf{R})\frac{Z_-}{\mathsf{R}} \quad (1.92)$$

$$G_{bxz}(\mathbf{R}\ ) = \frac{i}{4}H_2(\kappa\mathsf{R})\frac{X_-Z_-}{\mathsf{R}^2} \quad (1.93)$$

1.11 Depolarization Dyadic in 2D

In 2D the homogenous background particular solution is:

$$\mathbf{E}_s(\boldsymbol{\rho}) = g^2 \int_A \Delta\varepsilon \; d\boldsymbol{\rho}' \; \hat{\mathbf{G}}_b(\boldsymbol{\rho}, \boldsymbol{\rho}') \cdot \mathbf{E}(\boldsymbol{\rho}') = g^2 \int_A \Delta\varepsilon \; d\boldsymbol{\rho}' \; \hat{\mathbf{G}}_b(\mathsf{R}\;) \cdot \mathbf{E}(\boldsymbol{\rho}') \quad (1.94)$$

with $\mathsf{R} = X_- \mathbf{u}_x + Z_- \mathbf{u}_z$. The divergent part of the Green Tensor is isolated like in the 3D case:

$$\begin{aligned} \mathbf{E}_s(\boldsymbol{\rho}) &= g^2 \Delta\varepsilon \int_{A-A_\delta} d\mathbf{r}' \, \hat{\mathbf{G}}_b(\mathsf{R}\;) \cdot \mathbf{E}(\boldsymbol{\rho}') + \\ &+ \Delta\varepsilon \int_{A_\delta} d\mathbf{r}' \, \hat{\mathbf{G}}_b(\mathsf{R}\;) \cdot \mathbf{E}(\boldsymbol{\rho}), \end{aligned} \quad (1.95)$$

the last term becomes improper due to terms of the Green function that come from the Hankel functions $H_2(\kappa\mathsf{R})$. These in fact, when expanded in an infinitesimal area A_δ about the value $\mathsf{R} = 0$, become:

$$H_2(\kappa\mathsf{R}) \approx \frac{-4\mathtt{i}}{\pi(\mathsf{R}\kappa)^2}. \quad (1.96)$$

Hence as A_δ approaches zero and $\mathsf{R} << 1$:

$$\hat{\mathbf{G}}_b(\mathsf{R}\;) \simeq \frac{1}{2\pi} \frac{1}{\mathsf{g}^2} \frac{1}{\mathsf{R}^2} \; \mathbf{u}_\mathsf{R} \mathbf{u}_\mathsf{R} \quad (1.97)$$

where for convenience we have transformed the GT in cylindrical coordinates. A short-cut to making this transformation is to take the cartesian basis of last section, and set $X_- = \mathsf{R}$ and $Z_- = 0$.
We can now treat the improper integral with the divergence theorem like we have done in 3D:

$$\int_{A_\delta} d\boldsymbol{\rho} \;\; \hat{\mathbf{G}}_b(\boldsymbol{\rho}, \boldsymbol{\rho}') = -\frac{1}{2\pi} \frac{1}{g^2 \varepsilon_b} \int_{A_\delta} d\mathsf{R} \nabla \left(\frac{\mathbf{u}_\mathsf{R}}{\mathsf{R}} \right) = -\frac{1}{2\pi} \frac{1}{g^2 \varepsilon_b} \oint_{C_\delta} d\mathsf{R} \, \frac{\mathbf{n} \, \mathbf{u}_\mathsf{R}}{\mathsf{R}} \quad (1.98)$$

where C_δ is the curve that corresponds to the perimeter of the Area A and $\mathbf{n}$ is the normal to the surface. Hence we designate:

$$\hat{\mathbf{L}}(\boldsymbol{\rho}) \equiv \frac{1}{2\pi} \int_{C_\delta} dR \; \frac{\mathbf{u}_R \mathbf{n}}{R} = \int_{C_\delta} dR \; \frac{\mathbf{n} \, \mathbf{u}_R}{R} \quad (1.99)$$

and expressions of $\hat{\mathbf{L}}(\boldsymbol{\rho})$ in 2D are reported in [15]. In this thesis we only need $\hat{\mathbf{L}}(\boldsymbol{\rho})$, calculated at the center of a square mesh and this is: $\hat{\mathbf{L}} = 1/2(\mathbf{u}_x \mathbf{u}_x + \mathbf{u}_z \mathbf{u}_z)$.

$$\mathbf{E}_s(\boldsymbol{\rho}) = g^2 \, \Delta\varepsilon \int_{A-A_\delta} d\mathbf{r}' \hat{\mathbf{G}}_b(\boldsymbol{\rho}, \boldsymbol{\rho}') \cdot \mathbf{E}(\boldsymbol{\rho}') - \Delta\varepsilon \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \mathbf{E}(\boldsymbol{\rho}) \quad (1.100)$$

1.12 2D Discretized Lippmann-Schwinger and M

For a defect resting on either sides of an air-metal semi-space the 2D discretized Lippmann-Schwinger equation becomes:

$$\mathbf{E}(\boldsymbol{\rho}_i) = \mathbf{E}_b(\boldsymbol{\rho}_i) \; + \; g^2 \sum_{j\neq i} \hat{\mathbf{G}}_b(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j) \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_j) A_j + \tag{1.101}$$

$$+g^2 \, \hat{\mathbf{M}} \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_i) - \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_i) + \tag{1.102}$$

$$+g^2 \sum_j \hat{\mathbf{G}}_I(\boldsymbol{\rho}_i, \boldsymbol{\rho}_j) \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_j) A_j \tag{1.103}$$

where $A_i = A/N^2$. Now if we set R_i such that $\pi \mathsf{R}_i^2 = A_i$ we have:

$$\hat{\mathbf{M}}(\boldsymbol{\rho}_i) = \lim_{\delta A \to 0} \int_{A_i - \delta A} d\mathbf{r}' \; \hat{\mathbf{G}}_b(\boldsymbol{\rho}_i - \boldsymbol{\rho}') = \lim_{\mathsf{R}_0 \to 0} \int_0^{2\pi} d\theta \int_{\mathsf{R}_0}^{\mathsf{R}_i} d\mathsf{R}' \;\; \mathsf{R}' \; \hat{\mathbf{G}}(\mathsf{R}') \tag{1.104}$$

To derive the correction term $\hat{\mathbf{M}}$ in 2D, we need to go back to the expression of the Green Tensor for a homogeneous medium Eq.(1.88). Express all the factors $X_-/\mathsf{R} = \cos\alpha$; $Z_-/\mathsf{R} = \sin\alpha$ therein, through the polar angle α. Once we have done that, we should observe that, when we integrate Eq.(1.104) angularly over $\int_0^{2\pi} d\alpha$ the terms containing $\sin\alpha$, $\cos\alpha$, $\sin\alpha\cos\alpha$, $\cos\alpha^2 - 1/2$ and $\sin\alpha^2 - 1/2$ give zero. Hence we only need to consider those terms in Eq.(1.88) to Eq.(1.93) of the components G_{xx}, G_{yy}, G_{zz} that contain $H_0(\kappa\, r)$.
The Hankel function can be expanded for $\mathsf{R} \sim 0$

$$H_0(\kappa\, \mathsf{R}) \simeq \frac{2i}{\pi} \ln(\kappa\, \mathsf{R}) + \Delta \tag{1.105}$$

with $\Delta = \frac{2i}{\pi}(\varsigma - \ln 2) + 1$ and where ς is Euler's gamma.
Hence integrating over the angle gives:

$$\int_0^{2\pi} d\alpha \; G_{xx}(\mathsf{R}, \alpha) = \int_0^{2\pi} d\alpha \; G_{zz}(\mathsf{R}, \alpha) = \left(1 - \frac{\gamma^2}{2}\right) \; [\psi - \ln(\kappa \mathsf{R})] \tag{1.106}$$

and

$$\int_0^{2\pi} d\alpha \; G_{yy}(\mathsf{R}, \alpha) \;\; = \;\; \left(1 - \frac{k_y^2}{\mathsf{g}^2}\right) \; [\psi - \ln(\kappa \mathsf{R})] \tag{1.107}$$

where: $\nu_0 = ln(2) - \varsigma + \frac{1}{2} \approx 0.615931$ and $\psi = \nu_0 + \frac{i\pi}{2}$.
Now integrating radially and using the definite integrals:

$$\int_{\mathsf{R}_0}^{\mathsf{R}} d\mathsf{R} \;\; \mathsf{R} \; \psi = \frac{\mathsf{R}^2 - \mathsf{R}_0^2}{2} \; \psi \tag{1.108}$$

$$\int_{\mathsf{R}_0}^{\mathsf{R}} d\mathsf{R}\ \ \mathsf{R}\ \ln(\kappa\mathsf{R}) = \frac{1}{2}\left[\ln(\kappa\mathsf{R})\mathsf{R}^2 - \ln(\kappa\mathsf{R}_0)\mathsf{R}_0^2 - \frac{1}{2}(\mathsf{R}^2 - \mathsf{R}_0^2)\right] \tag{1.109}$$

we end up with:

$$M_{xx} = M_{zz} = \left(1 - \frac{\gamma^2}{2}\right)\frac{\mathsf{R}^2}{2}\left(\ \psi - \ln(\kappa\mathsf{R}) + \frac{1}{2}\ \right) \tag{1.110}$$

$$M_{yy} = \left(1 - \left(\frac{k_y}{\mathsf{g}}\right)^2\right)\frac{\mathsf{R}^2}{2}\left(\ \psi - \ln(\kappa\mathsf{R}) + \frac{1}{2}\ \right) \tag{1.111}$$

1.13 2D Green tensors of the Interface

It is trivial to convert the plane interface response tensor G_I from its 3D expression of Eq.(1.68) to its 2D version by carrying out the integration over y. Consider the indirect Green tensor for a source in the air semi-space $\hat{\mathbf{G}}_I^{(1,1)}(X_-, Z_+)$ with $Z_+ > 0$ and the indirect Green tensor for a source in the metal semi-space $\hat{\mathbf{G}}_I^{(2,2)}(X_-, Z_+)$ with $Z_+ < 0$

The spectral representation for the indirect Green Tensor in the air semi-space

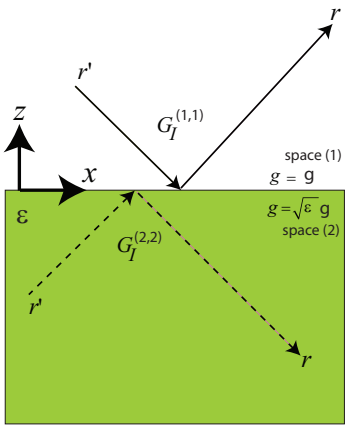

Figure 1.8: A representation of two kinds of indirect Green tensors in an air metal semi-space. Both tensors propagate light indirectly, that is propagates light by reflection at the air-metal interface. $\hat{\mathbf{G}}_I^{(1,1)}$ quantifies the light emission when both $\mathbf{r}$ and $\mathbf{r}'$ are the air semi-space, while $\hat{\mathbf{G}}_I^{(2,2)}$ quantifies the light emission when both $\mathbf{r}$ and $\mathbf{r}'$ are the metal semi-space.

is then:

$$\hat{\mathbf{G}}_I^{(1,1)}(X_-, Z_+) = \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty} dk_x\, e^{ik_x X_-}\,\frac{e^{ik_z Z_+}}{k_z}\,\left[r_p^{(1,1)}\mathbf{k}_p^-\mathbf{k}_p^+ + r_s^{(1,1)}\mathbf{k}_s\mathbf{k}_s\right] \tag{1.112}$$

Now let us write explicitly:

$$\mathbf{k}_p^-\mathbf{k}_p^+ = \frac{k_z^2}{g^2\,k_\parallel^2}\begin{pmatrix} {k_x}^2 & k_xk_y & k_xk_\parallel{}^2/k_z \\ k_xk_y & k_y^2 & k_yk_\parallel{}^2/k_z \\ -k_xk_\parallel^2/k_z & -{k_\parallel}^2k_y/k_z & -{k_\parallel}^4/k_z^2 \end{pmatrix} \tag{1.113}$$

and

$$\mathbf{k}_s\mathbf{k}_s = \frac{1}{k_\parallel}^2 \begin{pmatrix} k_y^2 & -k_x k_y & 0 \\ -k_x k_y & k^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.114}$$

keeping in mind $k_y = k_y^0$ is the y component of the incident wavevector.
Similarly the spectral representation for the indirect Green Tensor in the metal semi-space is then:

$$\hat{\mathbf{G}}_I^{(2,2)}(X_-, Z_+) = \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty} dk_x\, e^{ik_x X_-}\, \frac{e^{-ik_z^m Z_+}}{k_z^m}\left[r_p^{(2,2)}\mathbf{k}_p^{m+}\mathbf{k}_p^{m-} + r_s^{(2,2)}\mathbf{k}_s\mathbf{k}_s\right] \tag{1.115}$$

$$\mathbf{k}_p^{m+}\mathbf{k}_p^{m-} = \frac{(k_z^m)^2}{\varepsilon g^2\, k_\parallel^2} \begin{pmatrix} {k_x}^2 & k_x k_y & -k_x {k_\parallel}^2/k_z^m \\ k_x k_y & k_y^2 & -k_y {k_\parallel}^2/k_z^m \\ k_x k_\parallel^2/k_z^m & {k_\parallel}^2 k_y/k_z^m & -{k_\parallel}^4/(k_z^m)^2 \end{pmatrix} \tag{1.116}$$

In this second case we can find the reflection coefficients placing an incident field within the metal semi-space matching the electric field boundary conditions to the fields:

$$\mathbf{E}_p^{(2,2)}(z') = e^{ik_z^m z'}\mathbf{k}_p^{m+} + e^{-ik_z^m z'} r_p^{(2,2)}\mathbf{k}_p^{m-} \tag{1.117}$$

$$\mathbf{E}_p^{(2,1)}(z') = e^{ik_z z'}\, t_p^{(2,1)}\mathbf{k}_p^{+} \tag{1.118}$$

$$\mathbf{E}_s^{(2,2)}(z') = e^{ik_z^m z'}\mathbf{k}_s + e^{-ik_z^m z' z'} r_s^{(2,2)}\mathbf{k}_s \tag{1.119}$$

$$\mathbf{E}_s^{(1,2)}(z') = e^{ik_z z'}\, t_s^{(2,1)}\mathbf{k}_s \tag{1.120}$$

Comparing with the reflection coefficients at the other side of the interface we can yield the whole set relations between the Fresnel reflection coefficients at either side of the interface:

$$r_p^{(1,1)} = \frac{k_z^m - \varepsilon\, k_z}{k_z^m + \varepsilon\, k_z}, \qquad r_s^{(1,1)} = \frac{k_z - k_z^m}{k_z + k_z^m} \tag{1.121}$$

$$r_p^{(2,2)} = -r_p^{(1,1)}, \qquad r_s^{(2,2)} = -r_s^{(1,1)} \tag{1.122}$$

$$t_p^{(2,1)} = \frac{2k_z^m\sqrt{\varepsilon}}{k_z^m + \varepsilon\, k_z}, \qquad t_s^{(2,1)} = \frac{2k_z^m}{k_z^m + k_z} \tag{1.123}$$

$$t_p^{(1,2)} = \frac{2k_z\sqrt{\varepsilon}}{k_z^m + \varepsilon\, k_z}, \qquad t_s^{(1,2)} = \frac{2k_z}{k_z^m + k_z} \tag{1.124}$$

$$\frac{t_p^{(1,2)}}{k_z} = \frac{t_p^{(2,1)}}{k_z^m} \tag{1.125}$$

The relations just reported should be kept for reference throughout the treatment.

Chapter 2

Numerical Computation of Sommerfeld's Integrals

2.1 3D Sommerfeld Integrals

In the the near-field (the region of space close to the inhomogeneities) all Sommerfeld's integrals only have an analytical spectral representation. Attaining its spatial representation demands numerical integration of the spectral representation. In the 3D case for a source in the vacuum semi-space:

$$\hat{\mathbf{G}}_I(\mathbf{R}_\parallel, Z_+) = \frac{\mathtt{i}}{8\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{k}_\parallel \, e^{i\mathbf{k}_\parallel \cdot \mathbf{R}_\parallel} \, \frac{e^{ik_z Z_+}}{k_z} \, \left[r_p^{(1,1)} \mathbf{k}_p^- \mathbf{k}_p^+ + r_s^{(1,1)} \mathbf{k}_s \mathbf{k}_s \right]$$

In 3D we will only deal with $\hat{\mathbf{G}}_I^{(1,1)}(\mathbf{R}_\parallel, Z_+)$ thus from now on we we will omit the superscript. Remember $r_p^{(1,1)}$ and $r_s^{(1,1)}$ are defined in Eq.(1.121).
Before explaining how we conduct the numerical integration of this tensor one more simplification is due. Transforming the integral into cylindrical polar coordinates with $k_x = k_\parallel \cos\phi_k$, $k_y = k_\parallel \sin\phi_k$, $k_z = k_z$ we can integrate over ϕ_k so as to reduce the integration over k_x and k_y to integration over $k_\parallel$ only. Furthermore in so doing, we will change from the cartesian to the polar basis and this also has the benefit of reducing the number of non-zero components. This can be achieved using the Bessel's functions identities reported in the appendix. Integrating over ϕ_k transforms Eq.(2.1) into:

$$\hat{\mathbf{G}}_I(\mathbf{R}) = \frac{i}{4\pi g^2} \int_{-\infty}^{\infty} dk_\parallel \, k_\parallel \, \frac{e^{ik_z Z_+}}{k_z} \, \left[r_p^{(1,1)} \hat{\mathbf{\Upsilon}}_p(k_\parallel R_\parallel) + r_s^{(1,1)} \hat{\mathbf{\Upsilon}}_s(k_\parallel R_\parallel) \right] \quad (2.1)$$

with:

$$\hat{\boldsymbol{\Upsilon}}_p(k_\parallel R_\parallel) =$$

$$= \begin{pmatrix} k_z^2\left[J_0(k_\parallel R_\parallel) - J_2(k_\parallel R_\parallel)\right]/2 & 0 & ik_\parallel k_z J_1(k_\parallel R_\parallel) \\ 0 & k_z^2\left[J_0(k_\parallel R_\parallel) + J_2(k_\parallel R_\parallel)\right]/2 & 0 \\ -ik_\parallel k_z J_1(k_\parallel R_\parallel) & 0 & -{k_\parallel}^2 J_0(k_\parallel R_\parallel) \end{pmatrix}$$

$$\hat{\boldsymbol{\Upsilon}}_s(k_\parallel R_\parallel) = \begin{pmatrix} g^2\left[J_0(k_\parallel R_\parallel) - J_2(k_\parallel R_\parallel)\right]/2 & 0 & 0 \\ 0 & g^2\left[J_0(k_\parallel R_\parallel) + J_2(k_\parallel R_\parallel)\right]/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The basics of the numerical integration of this Sommerfeld Integral have been reported briefly in[9], while treatments and techniques are mentioned in[4, 31–33]. In this thesis we are going to provide a detailed account of the practical implementation of such 'standard' technique, explaining the problems we encountered and the solutions we have found. Furthermore we are going to present an innovative integration technique useful for large $R_\parallel$.

The numerical computation of Sommerfeld integrals can be time-consuming and nightmarish due to the presence of poles, branch-point and strongly oscillatory integrands. First of all, let us clarify a fundamental subdivision of the spectral domain of integration. The region $k_\parallel < g$ will be referred to as the propagating-modes region because any plane wave mode ($\mathbf{k}_p^\pm$ or $\mathbf{k}_s$) in this region corresponds to waves propagating in three dimensional space. Viceversa the argument $k_\parallel > g$ implies an exponential decay in the vertical direction (z), and so the waves are confined close to the plane surface. Consequently these are called evanescent modes since they only propagate between two points of the surface, but not away from it.

Secondly, we need to make some specifications about the integrand. In the propagating region, the integral in Eq.(2.1) is well-defined only if the analytical structure of the integrand is properly taken into account. Mathematically this is required because the integrand contains square roots. These terms produce a branch cut on the axis $Re(k_\parallel)$, so direct integration of Eq.(2.1) along this axis is in principle meaningless. For physically meaningful solutions the branch-cut and the radiation condition are constrained so that $e^{\mathrm{i}k_z Z}$ decays at infinity. We set $Z > 0$ in Chapter 1, which means that we must take $\Im\{k_z\} > 0$. A graphic representation of a branch-cut is provided by Fig.2.1

for the function $\sqrt{1-\xi^2}$. The function is ill-defined for $\xi < 1$ and crossing the branch-cut there, results in a discontinuous jump in the function's value. For $\xi > 1$ the function becomes continuous. The square roots contained in our integrand are $k_z = \sqrt{g^2 - {k_\parallel}^2}$ and thus the branch cut only concerns the propagating-mode region. The correct prescription for evaluating Eq.(2.1) in the propagating region can be obtained in three steps: a) Consider that the media in the radiation regions has infinitesimal absorption. This translates into a small imaginary part under the square roots (eg. $k_z = \sqrt{g^2 - {k_\parallel}^2 + i\eta}$, $k_z^m = \sqrt{\varepsilon g^2 - {k_\parallel}^2 + i\eta}$). b) Perform the integral for a finite η. c) Take the limit of the result as η tends to zero. The addition of η under the square roots (in step a), modifies the analytical structure of the integrand of Eq.(2.1): the branch cut is displaced away from the real $Re(k_\parallel)$ in the way represented in Fig. 2.4 . It is now possible to define the integral along this axis. Actually,

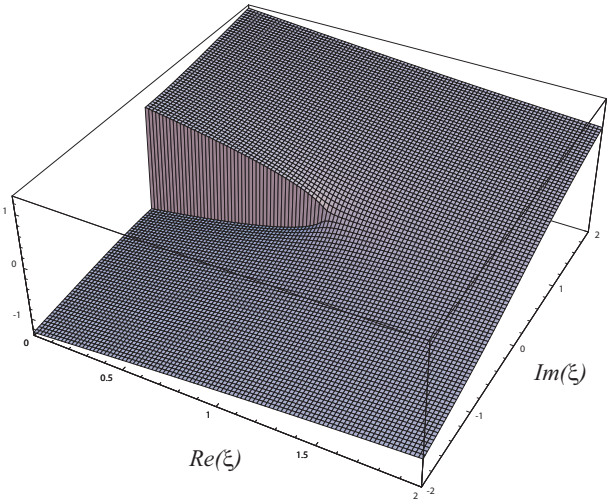

Figure 2.1: Schematic representation of the branch cut generated by the function $\sqrt{1-\xi^2}$, in the complex ξ plane.

the limit $\eta \to 0$ can be taken in the integrand, provided the cited branch cut is not crossed the integration path is deformed into the complex k plane. Accordingly we can use Cauchy's theorem to perform the integral on any path in the complex plane that does not cross the branch cut, as done in[9].
A subtle unexplained detail of this technique is how to convert an integral on the real axis into an integral over a complex curve. Practically we took the initial real variable $k_\parallel = k$ and add to it an imaginary part so that $k_\parallel = k + \mathtt{i}t(k)$. In so doing we can divert the integral from the real axis to an integral over a curve in the complex plane.

$$\int_0^c f(k)dk = \int_0^c f(k + \mathtt{i}\, t(k)) \left[1 + \mathtt{i}\, \frac{dk}{dt}\right] dk \tag{2.2}$$

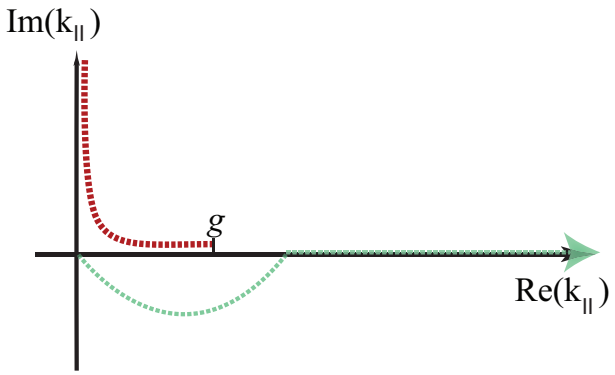

Figure 2.2: Schematic representation of the branch cut of k_z, in the complex $k_\parallel$ plane when the radiation regions have infinitesimal absorption. The figure also shows how integration on the real axis can be avoided in the region of the branch-cut, by using Cauchy's theorem and creating an equivalent integration path in the complex plane.

Our aim is to use this freedom to find an integration path with optimum computational speed. This is achieved studying the landscape of the integrand in the complex reciprocal space. Among other things, this reveals difficulties presented by the integrand is the the branch-point at $k_\parallel = g$, and remarkably the surface plasmon polariton pole in the Fresnel coefficients. Such pole corresponds to a particular evanescent mode which is indeed responsible for the long range transport of light between any two points of an interface separating a dielectric from a metal. This mode arises from the divergence of the Fresnel reflection coefficient $r_p^{(1,1)} = (k_z^m - \varepsilon\, k_z)/(k_z^m + \varepsilon\, k_z)$ which actually gives the response of the plane at each wavevector. Physically the divergence corresponds to a collective oscillation of the electron plasma in the surface, and the corresponding wave mode is denominated the surface plasmon polariton of the plane [34] and corresponds to the wavevector $k_\parallel = k_p$ defined by the condition $k_z^m = -\varepsilon k_z$. The singularity only regards the part of the integrand related to p-polarized waves. Both the branch-point singularity and the surface plasmon pole are represented in the complex reciprocal space in Fig.2.3.
Since the reflection coefficient $r_p^{(1,1)}$ is the only part of the integrand of Eq.(2.1) that depends on ε, if we choose an integration path that does not pass too close to neither the residue nor the branch point, such path will be very weakly dependent on ε. At least compared to the strong oscillations of the integrand, generated by the Bessel functions especially for large values of $R_\parallel$ and by the e^{ikzZ_+} term, in the propagating modes region, if Z_+ is large.
We shall now present our solution scheme. We have devised two integration techniques.
I. The first one for integrals in Eq.(2.1) with $R_\parallel < \lambda$ is named the standard integration technique because the main argument and ideas behind the solution

scheme have been sketched in the references cited. We shall explain in detail our implementation of the solution scheme, to highlight the solutions we found to facilitate the computation of Sommerfeld Integrals.

II. A second approach was devised for Sommerfeld Integrals whose space parameters have $R_{\|} > \lambda$. This was not reported in previous treatments but we found it became necessary to change the integration strategy, especially when $R_{\|}$ becomes several times larger than the free space wavelength.

Let us explain these techniques one by one.

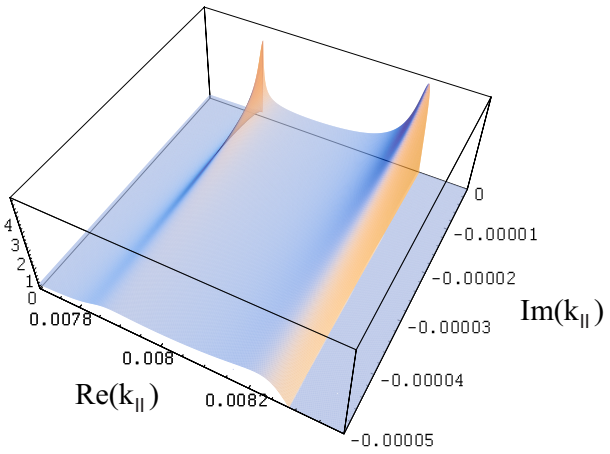

Figure 2.3: Schematic representation of the branch cut of k_z, in the complex k plane. $\Re\{Gzz\}$ for $\varepsilon = -10 + 0.3\mathtt{i}.$, $R_{\|} = 500nm$, $z = 50nm$ $\lambda = 800$

2.2 Standard Integration Technique $R_{\|} < \lambda$

2.2.1 Solution Scheme

The first path (1) of the solution scheme is the complex straight line $k_{\|} = k(1 - \mathtt{i}0.85)$ taken in the interval $k \in [0, g/2]$ which we traced and plotted in Fig.2.4, in Fig.2.5 and Fig.2.6. This first path is the first interval of the whole solution scheme rendered in the Fig.2.4.
The second paths (2) is another straight line: $k_{\|} = k - \mathtt{i}0.85g/2$ parallel to the $\Re\{k_{\|}\}$-axis, for the interval $k \in [g/2, g_e]$, with $g_e = 20\,\pi/(Z_+ + R_{\|})$
The third path (3) splits in two. The Bessel functions in the integrand of Eq.(2.1) are expressed as a sum of Hankel functions of the first and second kind. The part containing Hankel functions of the first kind, $H_n^{(1)}$, is integrated on the path $3(a)$: $g_e + \mathtt{i}t$ for $t \in [g/2, +\mathtt{i}10\pi/(Z_+ + R_{\|})]$, that is a straight line that runs parallel to the imaginary axis. The part containing Hankel functions of the second kind, $H_n^{(2)}$, is integrated on on the path $3(b)$: $g_e + \mathtt{i}t$ for $t \in [g/2, -\mathtt{i}10\pi/(Z_+ + R_{\|})]$, that is a straight line that runs parallel to the imaginary axis, but in the negative direction.

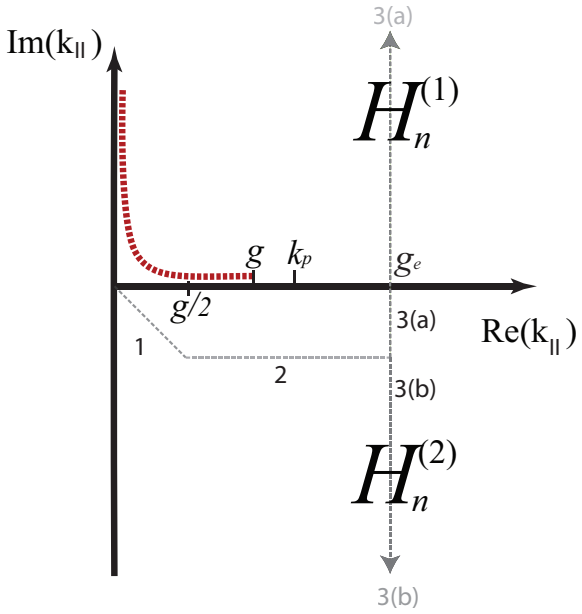

Figure 2.4: Schematic representation of the integration curve used. The curve consists of three paths(1-3). Path 3 is divided in two. The part of the integrand that contains Hankel functions of the first kind $H_n^{(1)}$ should be integrated in the positive $\Im\{k_{\|}\}$ direction, while the part of the integrand that contains Hankel functions of the second kind $H_n^{(2)}$ should be integrated in the negative $\Im\{k_{\|}\}$ direction.

2.2.2 The topology of Sommerfeld's Integrands

When integrating in the quadrant $t = \Im\{k_{\|}\} < 0$ over propagating modes we must be careful. On the one side we may find several advantages in diverting the integration to this region, basically we are damping down the pole and the branch-point, represented in Fig.2.3 in a close-up on the integrand in the complex $k_{\|}$ plane and, in doing so, the dependence of the complex path $t(k)$ on ε can be neglected.
On the other hand the integration strongly depends on the space coordinates $R_{\|}$ and Z_{+}. The oscillating functions $e^{ik_z Z_{+}}$ and $J_n(k_{\|}R_{\|})$ are superposed and built into the integrand. Therefore the behavior of the integrand in the complex reciprocal space is extremely hard to predict for arbitrary $R_{\|}$ and Z_{+}. For example the exponential factor in the complex reciprocal space becomes:

$$\exp\{\mathtt{i}\ Z_{+}\ \sqrt{g^2 - k^2 + t^2 - 2\mathtt{i}kt}\ \} \tag{2.3}$$

This can turn into a nicely integrable exponential decaying function or start to oscillate dramatically in different regions of the k-space. As to the Bessel functions, they can literally explode for not too large imaginary argument $t = \Im\{k_{\|}\}$. With this expression we intend to emphasize that the can reach infinity-type values such as $J_2(-100\mathtt{i}) \sim 10^{42} + \mathtt{i}\ 10^{35}$; see [35] for a detailed discussion. These paths are virtually inaccessible to numerical integration. Of course the explosion of the complex Bessel functions is progressively accelerated for larger $R_{\|}$ (and this point is particularly relevant towards understanding the alternative integration technique reported in the next section).
The situation is even more complex if we want our integration technique to continue to be efficient for large values of Z_{+} like in Fig.2.5. In this case we cannot keep the integration path to close to the $\Re\{k_{\|}\}$-axis in the propagating mode region, because the exponential in this region is a rapid sinusoidal oscillating function.
To sum up: we cannot integrate on the $\Re\{k_{\|}\} = k$-axis or in the positive $\Im\{k_{\|}\} = t$-axis and we can neither move too deep into large negative values of $\Im\{k_{\|}\} = t$-axis nor stay too close to small values of $\Im\{k_{\|}\} = t$-axis.

2.2.3 Path 1

We came up with a line where $t = -0.85k$ as a sort of paved lane to perform integration in the propagating region $k \in [0, g/2]$. The convergence of the integral dyadic was tested by executing the integrals by at least two independent integration methods. We found that large regions of the $(R_{\|}, Z_{+})$ space did not require cautious attentions; nevertheless we could always find some special values of $R_{\|}$ and Z_{+} where the integration did not easily converge. On a trial

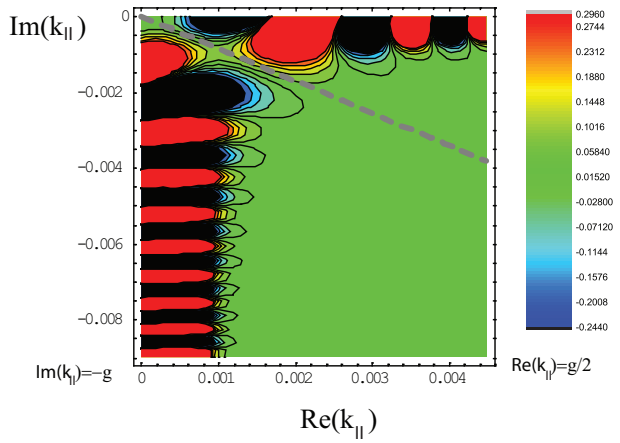

Figure 2.5: Contour Plot of $\Im\{G_{zz}\}$, $\varepsilon = -10$, $\lambda = 700nm$, $R_{\|} = 0.5\lambda$ $Z_+ = 20\lambda$, grey line $t = -0.85k$

and error basis and with the intent to dribble large amplitude oscillations exemplified in Fig.2.5, we observed that the $t = -0.85k$ line path heads straight to a region where the integrand settles down and that this was largely independent of ε.

We have actually determined this sort of "tunneling region" inductively. It

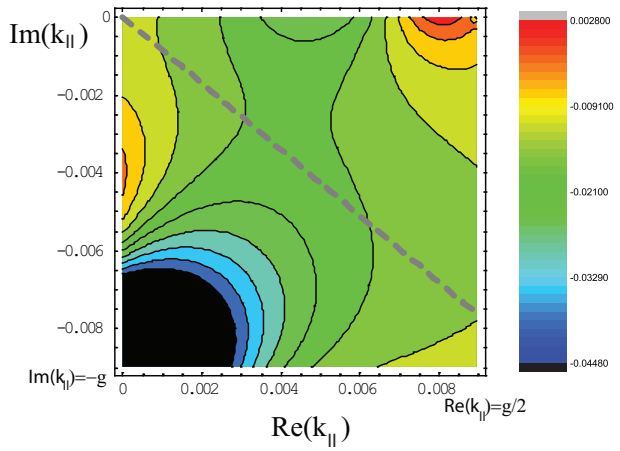

Figure 2.6: Contour Plot of $\Re\{G_{xx}\}$, $\varepsilon = -30$, $\lambda = 700nm$, $R_{\|} = 0.2\lambda$ $Z_+ = \lambda$, grey line $t = -0.85k$.

emerged naturally as a solution to problems of numerical integration but we can qualitatively explain why it works. First of all notice that the exponential of Eq.(2.1) and of Eq.(2.3) can be rewritten, in this region, as:

$$\exp\{-\Im\{\mathbf{k}_z\}Z_+\}\exp\{\mathtt{i}\Re\{\mathbf{k}_z\}Z_+\} \tag{2.4}$$

This is a standard damped oscillation. Now when, as in our case: $k < g$ the sinusoidal factor $\exp\{\mathtt{i}\Re\{\mathbf{k}_z\}Z_+\}$ is connected to the argument $g^2-k^2+t^2$ under the square root of Eq.(2.3). Conversely the damping factor $\exp\{-\Im\{\mathbf{k}_z\}Z_+\}$ depends on the $2\mathtt{i}kt$ argument under the square root of Eq.(2.3). Having clarified this we can deduce first of all that close to either the imaginary axis or the real axis (where kt tends to zero) the exponential is a sinusoidal function, and secondly that the integrand is mostly damped at the center of the quadrant where argument $t \simeq -k$ (we actually found $t = -0.85k$ is best). Finally the path $t = -0.85k$ for integration is only used up to the value $k = g/2$, to prevent the Bessel functions in the integrand from having a large imaginary argument.
In conclusion the remedy value provided particularly by the first segment of the path are evident in Fig.2.5 as applied to some of the most problematic integrands.

2.2.4 Path 2

Now let us move on to explaining the rest of the integration curve. In Fig.2.7 we show the path (1) and link it to the path(2). The path 2 is consequential

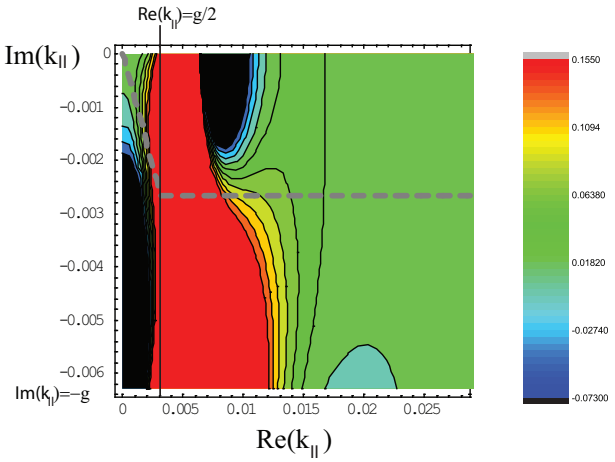

Figure 2.7: Contour Plot of $\Re\{G_{zz}\}$, $\varepsilon = -40$, $\lambda = 1000nm$, $R_{||} = 0.2\lambda$ $Z_+ = 0.4\lambda$, grey line $t = -0.85k$ and parallel segment

to the choice of the first segment. Once we terminated the integration over the first path: $\Re\{k_{||}\} = g/2$ and we are at the point $k_{||} = g/2 - \mathtt{i}0.85g/2$ see Fig.2.4 and Fig.2.7. Hence we need to finish up integration from $k = g/2$ to infinity. The choice we made is to continue the integration parallel to the $\Re\{k_{||}\}$ axis. Such choice was straightforward because we are in-between the Bessel functions oscillations and the branch-point and residue. Hence we integrate

past $\Re\{k_\parallel\} = g$ on a path $k_{\parallel_e} = k - \mathtt{i}0.85g/2$ parallel to the $\Re\{k_\parallel\}$-axis, for an interval: $k \in [g/2, g_e]$. Where we have located g_e past the plasmon residue $g_e > k_p$, so that this path takes us to the integration of the evanescent modes. This second path ends at the point in k-space $k_{\parallel_e} = g_e/2 - \mathtt{i}0.85g/2$. To establish the value of g_e we set g_e one order of magnitude greater than a wavevector value obtained by the direct real space vector $g_e = 20\pi/(Z_+ + R_\parallel)$. We defined such point g_e, to make the technique valid for the largest number of $R_\parallel$ and Z_+ and at the same time ensure that is safely far enough from the singularities.

2.2.5 Path 3

The integration technique was set off to substitute integration over the real axis with that over a complex curve by using the Cauchy theorem. This means we need to return from path(2) to the real axis and integrate up to infinity on the real axis, like in Fig.2.2. This would make the entire integral on the complex curve equivalent to the one on the real axis, by Cauchy's theorem. In fact the plasmon residue is displaced by adding infinitesimal absorption to the integrand through the same prescription we have described for the branch-cut. It is thus to be considered as located in the 0^+ quadrant and in the closed path we have chosen the residue is not encircled.
However from path(2) we can also find a third path that is in turn equivalent to the integration on the real axis, still using Cauchy's theorem. This is relevant because in many cases integrating the real axis is inconvenient and inefficient. In fact the Sommerfeld integral can exhibit a long oscillatory tail coming from the Bessel functions on the real axis (with a real argument) and the exponential factor. The latter for $k > g$ is well defined and also turns into a damped oscillation $\exp\{-\Im\{\mathbf{k}_z\}Z_+\}\exp\{\mathtt{i}\Re\{\mathbf{k}_z\}Z_+\}$ which is difficult to integrate, especially for small Z_+. The technique to handle the long oscillatory tail of the Sommerfeld integral is prescribed in [9]. First we are going to detail the general technique and after that, we are going to explain how we have applied it to yield the two-way path in Fig.2.4.
According to the general prescription we represent the Bessel functions as a sum of Hankel functions of first and second kind $J_n = H_n^{(1)} + H_n^{(2)}$ then we perform the integration of each Hankel function on a different path. As shown in Fig.2.8, we can replace the integration of an integrand containing a Bessel function over the interval of the real axis $[g_e, \infty]$ by closing the integration path in the upper and lower quadrant and integrating respectively the Hankel functions of the first and second kind. With reference to the figure let us define the following curves:
a) $\Gamma_1 = k$ for $k \in [g_e, +\infty]$ is the straight line that runs on the real axis from

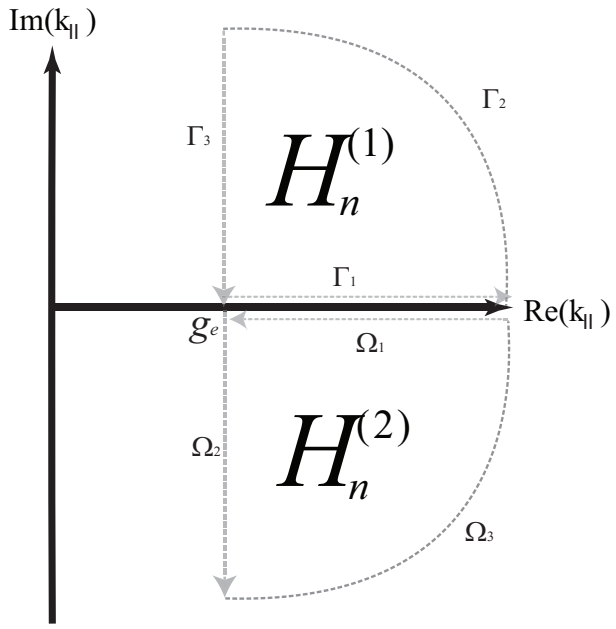

Figure 2.8: Two closed path in the complex reciprocal space of $k_{\|}$: $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$ and $\Omega = \Omega_1 + \Omega_2 + \Omega_3$. The Hankel function of the first kind $H_n^{(1)}(z)$ decays at infinity when $\Im\{k_{\|}\} > 0$, viceversa the Hankel function of the second kind $H_n^{(2)}(z)$ decays at infinity when $\Im\{k_{\|}\} < 0$.

g_e onwards.
b) $\Gamma_3 = g_e + \mathtt{i}t$ for $t \in [+\mathtt{i}\infty, 0]$ is the straight line that runs parallel to the imaginary axis, but in the negative direction. It comes from complex infinity and stops on the $\Re\{k_{\|}\}$-axis at g_e.
c) Γ_2 is any curve (for example a quarter of a circle) that joins Γ_1 to Γ_3, generating a closed path $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$.
Likewise, see Fig.2.8, another closed path is generated by $\Omega = \Omega_1 + \Omega_2 + \Omega_3$.
d) $\Omega_1 = -\Gamma_1 = -k$ for $k \in [+\infty, g_e]$ is the straight line that runs on the real axis and in the negative direction from infinity to g_e .
e) $\Omega_3 = g_e + \mathtt{i}t$ for $t \in [0, -\mathtt{i}\infty]$ is the straight line that runs parallel to the imaginary axis but in the negative direction, and crosses the $\Re\{k_{\|}\}$-axis at g_e.
f) Ω_2 is any curve (for example a quarter of a circle) that joins Ω_1 to Ω_3.
By definition of a Bessel function we have that :

$$\int_{\Gamma_1} f(k_{\|})\, J_n(k_{\|}R_{\|})\, dk_{\|} = \frac{1}{2}\int_{\Gamma_1} f(k_{\|})\, H_n^{(1)}(k_{\|}R_{\|})\, dk_{\|} + \\ + \frac{1}{2}\int_{\Gamma_1} f(k_{\|}) H_n^{(2)}(k_{\|}R_{\|})\, dk_{\|} \qquad (2.5)$$

where the closed path is meant to be taken arbitrary large in size, as the curves Ω_2, Γ_3 actually span from 0 to infinity. However using the fact that

$H_n^{(1)}(z) \sim e^{iz}$ and $H_n^{(2)}(z) \sim e^{-iz}$ the Hankel function vanish respectively at $+\infty$ and $-\infty$ and hence:

$$\int_{\Gamma_2} dk_\parallel f(k_\parallel) H_n^{(1)}(k_\parallel R_\parallel) = 0 \tag{2.6}$$

$$\int_{\Omega_3} dk_\parallel f(k_\parallel) H_n^{(2)}(k_\parallel R_\parallel) = 0 \tag{2.7}$$

Therefore applying Cauchy's theorem to these closed paths we know that:

$$\oint_{\Gamma} dk_\parallel f(k_\parallel) H_n^{(1)}(k_\parallel R_\parallel) = 0 \tag{2.8}$$

$$\oint_{\Omega} dk_\parallel f(k_\parallel) H_n^{(2)}(k_\parallel R_\parallel) = 0 \tag{2.9}$$

If we set $\Gamma' = -\Gamma_3$ means that we can integrate the term $H_n^{(1)}(z)$, in Eq.(2.5), on the path Γ' while the term containing $H_n^{(2)}(z)$, over the path Ω_2 and Eq.(2.5) becomes:

$$\begin{aligned}\int_{\Gamma_1} f(k_\parallel) J_n(k_\parallel R_\parallel) dk_\parallel = &+\frac{1}{2}\int_{\Gamma'} f(k_\parallel) H_n^{(1)}(k_\parallel R_\parallel) dk_\parallel \\ &+\frac{1}{2}\int_{\Omega_2} f(k_\parallel) H_n^{(2)}(k_\parallel R_\parallel) dk_\parallel.\end{aligned} \tag{2.10}$$

To sum up, applying the Cauchy theorem to the last bit of integration on the real axis $[g_e, \infty]$, means we can replace the real axis integration of a component of the integrand containing a Bessel function, with a two paths integrating a Hankel function on the imaginary axis. The constraint is, however, that $H_n^{(1)}$ is integrated in the positive imaginary direction and $H_n^{(2)}$ in the negative imaginary direction.

Now we are going to explain how we have adapted this known technique for handling the oscillatory tail of the Sommerfeld Integral to our integration path. At the end of the second integration segment we are at the point of complex space $g_e + \mathtt{i}g/2$. The technique substitutes Bessel function integration into Hankel functions integration, from the point on the real axis g_e. Therefore to apply the technique we need to take into account the integration over a path Δ connecting the point $g_e + \mathtt{i}g/2$ to the real axis at the point g_e. This is rendered in Fig.2.9. Hence to finish up our integration of the Sommerfeld integral from the end of the second path we need to evaluate the integral over the path $\Delta + \Gamma_1$. Now using the equivalence Eq.(2.10) between the integration

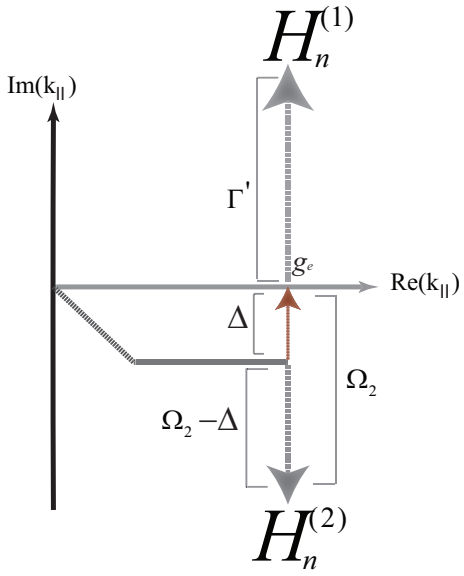

Figure 2.9: The two-way path of the standard integration technique is represented. In path 3 we integrate the part of the integrand containing $H_n^{(1)}$ on $\Delta + \Gamma'$ and the part of the integrand containing $H_n^{(2)}$ on $\Omega_2 - \Delta$.

over the path Γ_1, we end up with:

$$\int_{\Gamma_1+\Delta} dk_{\|} f(k_{\|}) J_n(k_{\|} R_{\|}) = \frac{1}{2} \int_{\Gamma'+\Delta} dk_{\|} f(k_{\|}) H_n^{(1)}(k_{\|} R_{\|}) + \\ + \frac{1}{2} \int_{\Omega_2-\Delta} dk_{\|} f(k_{\|}) H_n^{(2)}(k_{\|} R_{\|}). \qquad (2.11)$$

This means that, from the end of the second integration segment at the point $g_e + \mathtt{i}g/2$, we can split the integration of the Bessel function in Hankel function integration on the appropriate direction of the imaginary axis, directly as rendered in Fig.2.4, and this is actually taking into account the integration over Δ. Once again how long the integration over the imaginary axis has to be carried on, is a function of the spatial parameters. We found that integrating from $g_e + \mathtt{i}g/2$ up to a value of $g_e \pm \mathtt{i}10\pi/(Z_+ + R_{\|})$ (in each direction) will guarantee the converged value of the integral, with a precision of two significant digits.

2.3 Modified Integration technique for large values of $R_{||}$

To the best of our knowledge the technique we are going to explain in this section is not described in the literature.
When $R_{||} > \lambda$ we find that that the standard integration procedure of the last section must be modified. A different approach is necessary to speed up the computation of the Sommerfeld integral when $R_{||}$ is slightly greater than the free-space wavelength and to avoid numerical problems when $R_{||} >> \lambda$. The reason for this is that, as observed in the last section, the Bessel functions of imaginary integrand cause huge divergences. Thus, when $R_{||}$ is several times larger then the free space wavelength, all Bessel functions will in general have a large argument, and the addition of a small imaginary part $t = \Im\{k_{||}\}$ translates into a very large imaginary argument in the Bessel function.
This occurs for $k < g$, for which the branch-cut prevents us from diverting

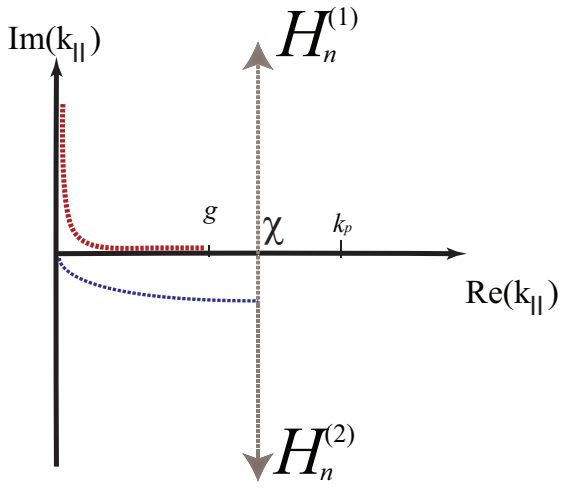

Figure 2.10: Schematic representation of the integration curve used in the Modified Integration technique for large values of $R_{||}$. Like in the standard technique, the curve consists of three paths(1-3). Path 3 is divided in two. The first path crosses the real axis, between $k_{||} = g$ and $k_{||} = k_p$. The part of the integrand that contains Hankel functions of the first kind $H_n^{(1)}$ should be integrated in the positive $\Im\{k_{||}\}$ direction, while the part of the integrand that contains Hankel functions of the second kind $H_n^{(2)}$ should be integrated in the negative $\Im\{k_{||}\}$ direction.

the integration into the semi-plane $\Im\{k_{||}\} > 0$. Thus we must use a different integration scheme. Remember that we need to integrate in the region of the k-space where $\Im\{k_{||}\} < 0$ at least over propagating modes $k_{||} < g$, to avoid the branch cut. Hence basically we keep as close as possible to the real axes

to a value where the imaginary argument of the Bessel function always gives $J(k_{\|}R_{\|}) \sim 1$. This was done by choosing a path:

$$k_{\|} = k - \frac{\mathtt{i}}{R_{\|}} \sin\left(\frac{\pi k}{2\chi}\right) \tag{2.12}$$

for $k \in [0, \chi]$.
Integrating evanescent modes for very large $R_{\|}$ is however daunting in the region of k-space where $\Im\{k_{\|}\} < 0$ and we resort to the integration of Hankel functions over the imaginary axes as soon as we get past $k_{\|} = g$. Hence this second technique consists in setting $\chi = g + 0.2(\Re\{k_p\} - g)$ as in Fig.2.10. This point is chosen to tunnel in-between the branch point and the residue. However we found it was best to stay closer to the branch-point than to the singularity. In fact the divergence of integrand at the branch-point goes like $1/\sqrt{k-g}$, while at the singularity it goes like $1/(k-k_p)$.
Notice χ depends on ε through k_p, as will become evident below.
Hence we have modified the standard technique by replacing $g_e > k_p$ with $\chi < k_p$. In this case we actually need to extract its contribution analytically from the Sommerfeld Integral. This is because when we apply the Hankel integration procedure, in Fig.2.8, we are encircling the residue at $k = k_p$. In fact from the residue theorem we have then:

$$\oint_{\Gamma} dk_{\|} f(k_{\|}) H_n^{(1)}(k_{\|}R_{\|}) \quad = \quad 2\pi\mathtt{i}\ Res\{f(k_{\|}) H_n^{(1)}(k_{\|}R_{\|})\}, \tag{2.13}$$

where the curve $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$ and where Γ_1,Γ_2 and Γ_3 are defined as in the last section but with $g_e = \chi$.
Extracting the residue is convenient for large $R_{\|}$, because this residue is an analytical function and takes care of a large contribution to the net Green Tensor in general, but most especially for points close to the surface $Z_+ \sim 0$ where the remaining integral turns out to be just a correction to the residue value.
To calculate the residue start from its definition. First of all determine the value of k_p in terms of ε and the free space wavevector:

$$k_{pz}^m + \varepsilon\, k = 0 \Longleftrightarrow k_p = g\sqrt{\frac{\varepsilon}{\varepsilon+1}} \quad k_{pz} = \frac{\mathtt{i}g}{\sqrt{-\varepsilon-1}} \tag{2.14}$$

Hence the value of the residue is obtained by taking the limit:

$$Res\{f(k_{\|}) H_n^{(1)}(k_{\|}R_{\|})\} = \lim_{k_{\|}\to k_p} (k_{\|} - k_p) f(k_{\|}) H_n^{(1)}(k_{\|}R_{\|}) \tag{2.15}$$

and since:

$$\lim_{k_{\|}\to k_p} \frac{(k_{\|} - k_p)}{k_z^m + \varepsilon\, k_z} = \lim_{k_{\|}\to k_p} \frac{k_z^m + \varepsilon\, k_z}{(\varepsilon^2-1)(k_{\|} + k_p)} = \frac{k_{pz}^m}{k_p(\varepsilon^2-1)} \tag{2.16}$$

the residue is:

$$\hat{\mathbf{G}}_p(\mathbf{R}) = 2\pi \mathtt{i}\ Res\{f(k_\parallel)H_n^{(1)}(k_\parallel R_\parallel)\} = -\frac{k_{pz}\,\varepsilon}{\varepsilon^2-1}\hat{\boldsymbol{\Upsilon}}_p(k_p R_\parallel)\ e^{ik_{pz}(z+z')} \quad (2.17)$$

$$\begin{aligned}\hat{\boldsymbol{\Upsilon}}_p(k_p R_\parallel) &= k_{pz}^2\frac{H_0^{(1)}(k_pR_\parallel)-H_2^{(1)}(k_pR_\parallel)}{4}\mathbf{u}_{R_\parallel}\mathbf{u}_{R_\parallel} + \\ &+ \mathtt{i}k_pk_{pz}\frac{H_1^{(1)}(k_pR_\parallel)}{2}(\mathbf{u}_{R_\parallel}\mathbf{u}_\phi - \mathbf{u}_\phi\mathbf{u}_{R_\parallel}) + \\ &+ k_{pz}^2\frac{H_0^{(1)}(k_pR_\parallel)+H_2^{(1)}(k_pR_\parallel)}{4}\mathbf{u}_\phi\mathbf{u}_\phi - k_p^2\frac{H_0^{(1)}(k_pR_\parallel)}{2}\mathbf{u}_z\mathbf{u}_z\end{aligned} \quad (2.18)$$

This technique sped up the calculations for large values of $R_\parallel$ with respect to the technique adopted for $R_\parallel < \lambda$. Indeed we have compared the numerical calculation of the Sommerfeld Integral with the two techniques. For example for a plane $\varepsilon = -10$ at free space wavelength $\lambda = 900nm$ and distances $R_\parallel = 4\lambda$ and $Z_+ = 0.1\lambda$ the modified technique only needs 70 points to get a converged value with a relative error of less that 10%, while applying the standard technique this requires 9000 points.
If we change $Z_+ = 0.01\lambda$ and $\varepsilon = -30$, but keep all the rest of parameters fixed, we need 90 points with the modified technique while 7000 points are needed with the standard technique. For $\lambda = 900nm$, $R_\parallel = 10\lambda$, $Z_+ = 0.001\lambda$ and $\varepsilon = -15$, is one of the cases in which the integral is just a small correction to the analytical value of the residue and we can attain this correction through the modified technique with 70 points. With the standard technique the integral keeps failing to converge with as much as 5×10^5 points.
However for $R_\parallel \sim \lambda$ the two techniques have comparable efficiency. For this reason we have chosen the value of $R_\parallel = \lambda$ as the value at which our computer code switches from the standard to the modified technique. For instance in the previous example if we change to $R_\parallel = 2\lambda$ and keep the rest of the parameters fixed we need 50 points with the modified technique and 500 with the standard technique.

2.4 2D Sommerfeld Integrals

In this treatment we are going to study 2D defects located either above a metal plane and below the metal plane interface, defined in Sec.1.13. Consequently we needed to evaluate two Sommerfeld Integrals namely Eq.(1.112) and Eq.(1.115). The 2D sommerfeld integrals integration is obtained with the same techniques just described for both of them. The Standard method is employed for $X_- < \lambda$ and the Modified technique is employed for $X_- > \lambda$. However a few differences pop up in the 2D integration and we shall now address them.
First of all notice that this time we are integrating in cartesian coordinates and this brings about some changes. The horizontal distance X_- can be either positive or negative. Nonetheless thanks to the symmetry of the Green tensor parallel to the surface we just need to evaluate the integral for one of the two, for instance take $X_- > 0$. Consider now the Sommerfeld integral in the air semi-space is then: $X' = -X_-$

$$\hat{\mathbf{G}}_I^{(1,1)}(X', Z_+) = \frac{\mathtt{i}}{4\pi} \int_{-\infty}^{\infty} dk_x \, e^{ik_x X'} \, \frac{e^{ik_z Z_+}}{k_z} \, \left[r_p \, \mathbf{k}_p^- \mathbf{k}_p^+ + r_s \, \mathbf{k}_s \mathbf{k}_s \right] \tag{2.19}$$

$$\mathbf{k}_p^- \mathbf{k}_p^+ = \frac{k_z^2}{g^2 \, k_\parallel^2} \begin{pmatrix} {k_x}^2 & k_x k_y & k_x {k_\parallel}^2 / k_z \\ k_x k_y & k_y^2 & k_\parallel^2 k_y / k_z \\ -k_x k_\parallel^2 / k_z & -{k_\parallel}^2 k_y / k_z & -{k_\parallel}^4 / k_z^2 \end{pmatrix} \tag{2.20}$$

and

$$\mathbf{k}_s \mathbf{k}_s = \frac{1}{{k_\parallel}^2} \begin{pmatrix} k_y^2 & -k_x k_y & 0 \\ -k_x k_y & k^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.21}$$

we have the property that the components that contain odd powers of k_x are antisymmetric with respect to X_- eg. $G_I(-X_-, Z_+) = -G_I(X_-, Z_+)$ while those that contain even powers of k_x are symmetric with respect to X_-. Thus the components G_{sxx}, G_{syy}, G_{szz}, G_{syz} and G_{szy} are even, while the components, G_{sxy}, G_{sxz}, G_{syx} and G_{szx} are odd. The same goes for $G_s^{(2,2)}$.
One further difference from the 3D integration is that in 2D the integration in k_x has to be actually carried out from $-\infty$ to $+\infty$ and thus we actually have two branch-cuts. These are located as in Fig.2.11, and can actually be avoided by mirroring the integration path that we used in 3D from 0 to g_e with respect

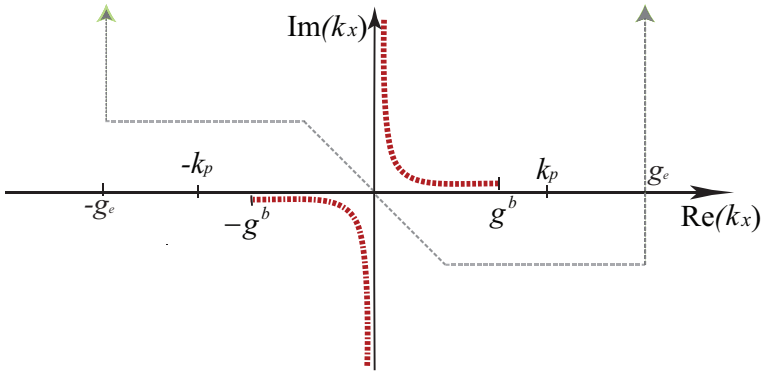

Figure 2.11: Schematic representation of the solution scheme for 2D sommerfeld integrals. The branch cuts of k_z, in the complex k_x plane are also represented when the propagating mode regions have infinitesimal absorption $g^b = \sqrt{g^2 - k_y^2 + \mathtt{i}\eta}$.

to the $\Re(k_x)$ axis. This time however we do not have Bessel functions in the integrand but instead we have the exponential factor $e^{ik_x X}$. Since we picked $X_- > 0$ we can use the same technique of integrating over the imaginary axes to avoid the long oscillatory tail of the Sommerfeld integral, but this time we always have to integrate in the positive direction of the imaginary axis. In fact this way we get that $e^{ik_x X}$ becomes a decaying exponential function that goes to zero at infinity $|k_x|$, and only then the Cauchy theorem is applicable.
The last difference is in the calculation of the residue that we use in the Modified integration scheme for $X_- > \lambda$. In the 2D case the residue is actually found at : $k_{px} = \sqrt{k_p^2 - k_y^2} = \sqrt{\frac{\varepsilon}{\varepsilon+1} g^2 - k_y^2}$ hence when taking the limit we get a slightly different result:

$$\lim_{k_x \to k_{px}} \frac{(k_x - k_{px})}{k_z^m + \varepsilon\, k_z} = \lim_{k_x \to k_{px}} \frac{k_z^m + \varepsilon\, k_z}{(\varepsilon^2 - 1)(k_x + k_{px})} = \frac{k_{pz}^m}{k_{px}(\varepsilon^2 - 1)} \tag{2.22}$$

Finally the surface plasmon residue contribution in 2D for both Sommerfeld integrals, defined in Sec.1.13, can be written as:

$$G = \frac{-\varepsilon^2}{\varepsilon^2 - 1} \frac{k_{pz}^3}{k_p^2\, g^2 k_{px}} \tag{2.23}$$

$$\hat{\mathbf{G}}_p^{(1,1)}(X_-, z, z') = G\, e^{ik_{px}X_-} e^{ik_{pz}Z_+}\ \hat{\mathbf{\Gamma}}_{spp}^{(1,1)} \tag{2.24}$$

$$\hat{\mathbf{G}}_p^{(2,2)}(X_-, z, z') = G\, e^{ik_{px}X_-} e^{-ik_{pz}^m Z_+}\ \hat{\mathbf{\Gamma}}_{spp}^{(2,2)} \tag{2.25}$$

$$\hat{\mathbf{\Gamma}}_{spp}^{(1,1)} = \begin{pmatrix} {k_{px}}^2 & k_{px}k_y & k_{px}{k_p}^2/k_{pz} \\ k_{px}k_y & k_y^2 & k_y{k_p}^2/k_{pz} \\ -k_{px}k_p^2/k_{pz} & -{k_p}^2k_y/k_{pz} & -{k_p}^4/k_{pz}^2 \end{pmatrix} \tag{2.26}$$

$$\hat{\mathbf{\Gamma}}_{spp}^{(2,2)} = \begin{pmatrix} {k_{px}}^2 & k_{px}k_y & -k_{px}{k_p}^2/k_{pz} \\ k_{px}k_y & k_y^2 & -k_y{k_p}^2/k_{pz} \\ k_{px}k_p^2/k_{pz} & {k_p}^2k_y/k_{pz} & -{k_p}^4/k_{pz}^2 \end{pmatrix}. \tag{2.27}$$

As explained in Sec.1.9, for a $2D$ system k_y is cyclic and fixed by the value established by the incident wave.

Chapter 3

Asymptotic Expressions for the 2D Green tensor in the far-field

3.1 The field outside the region of the defect

So far we have reported on how we find self-consistently the field *inside the source region* or the region (Area or Volume) of the defect. If the background has dielectric constant ε_b and the defect ε, the field within the defect is the solution to this Maxwell equation:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}') - \varepsilon\, g^2\, \mathbf{E}(\mathbf{r}') = 0, \quad \mathbf{r}' \in V \tag{3.1}$$

In this section we are going to be concerned with finding the field emitted by the source or scatterer, *outside the source region*. Notice that such field is the

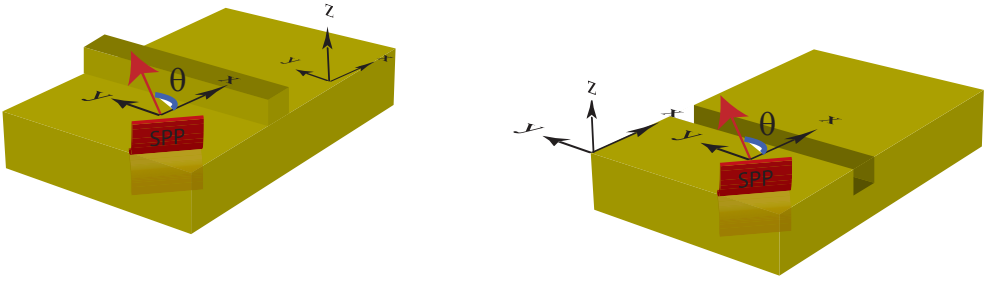

Figure 3.1: A schematic representation of 2D surface defects. These are defects that are infinite in the y-direction and can be either protrusions or indentations. 2D surface defects scatter a surface plasmons, impinging at an angle θ.

solution to a different Maxwell equation:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - \varepsilon_b\, g^2\, \mathbf{E}(\mathbf{r}) = 0, \quad \mathbf{r} \in \Re^3/\{V\} \tag{3.2}$$

Unless stated otherwise, in this chapter *we are only going to be concerned with 2D system* (as in Fig.3.1), where V is replaced for A. Remember that all fields exhibit this functional form:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}(\boldsymbol{\rho})\, e^{ik_y y}. \tag{3.3}$$

Hence as a shorthand notation we simply refer to the field $\mathbf{E}(\mathbf{r})$ as $\mathbf{E}(\boldsymbol{\rho})$, where $\mathbf{r} = x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z$ while $\boldsymbol{\rho} = x\mathbf{u}_x + z\mathbf{u}_z$. **By doing so we consider the y-dependence implicit in our equations**.
In this case the solution outside the source region is:

$$\mathbf{E}(\boldsymbol{\rho}) = \mathbf{E}_b(\boldsymbol{\rho}) + g^2\,\Delta\varepsilon \int_A d\boldsymbol{\rho}'\hat{\mathbf{G}}(\boldsymbol{\rho},\boldsymbol{\rho}')\cdot\mathbf{E}(\boldsymbol{\rho}') \tag{3.4}$$

where $\mathbf{E}_b(\boldsymbol{\rho})$ is the back-ground field defined in Chapter 1.
This is the solution for the net field everywhere outside A: $\boldsymbol{\rho} \in \Re^2/\{A\}$. Eq.(3.4) looks like the solution in the source region (which we designate the Lippmann-Schwinger equation) but is a different equation, that is valid for the field outside A. The main difference between the Eq.(3.4) and Lippman-Schwinger equations, apart from the absence of the depolarization tensor, is that in Eq.(3.4) the variables $\boldsymbol{\rho}'$ and $\boldsymbol{\rho}$ are completely decoupled. In fact the field at $\boldsymbol{\rho}'$ inside A, acts as a source for the field at $\boldsymbol{\rho}$ outside of A, but not the other-way round (as it happens when both $\boldsymbol{\rho}'$ and $\boldsymbol{\rho}$ are located within A).
One of the main advantages of the Green tensor is that, far from the source, we can always get an analytical expression for the Sommerfeld Integrals or Green tensors $\hat{\mathbf{G}}_I$. This allows us to solve Eq.(3.4) once we know the source field. This in turn gives us a chance to associate scattering problems with the analytical solutions for the emission by equivalent dipole distributions.

3.2 The Transmitted 2D Green Tensor

As said we are only going to be considering a background semi-space filled with air (space 1, z>0) and metal (space 2, z<0).

First of all we are going to resolve how the emission of surface plasmon

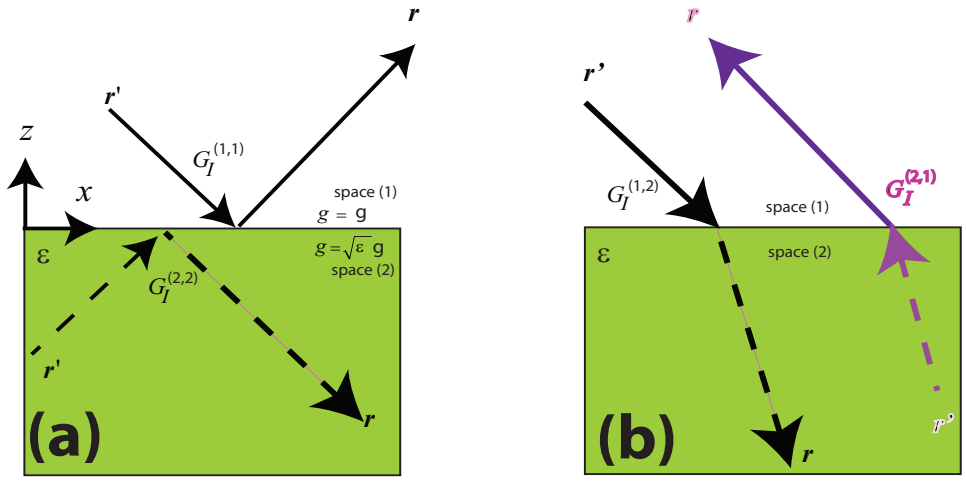

Figure 3.2: An explanatory representation of the indirect Green tensors in an air-metal semi-space. The indirect Green tensor $\hat{\mathbf{G}}^{1,1}$ and $\hat{\mathbf{G}}^{2,2}$ propagate light from a source to a point in the same semi-space by reflection at the interface. On the other hand $\hat{\mathbf{G}}^{(1,2)}$ and $\hat{\mathbf{G}}^{(2,1)}$ propagate light from a source in one semi-space to a point in the other semi-space by tansmission at the interface.

polariton by a metal surface stems directly from the Green Tensor formalism. So far we have only developed the Green tensor relevant to calculating the field in the source region A (in 3D it was V). With the exception of Sec.3.4.3 in this chapter we will be treating bi-dimensional systems. The bi-dimensional Green tensors we have already seen, in Sec.2.4, are: $\hat{\mathbf{G}}^{(1,1)}$ if the defect or source is above the metal interface and $\hat{\mathbf{G}}^{(2,2)}$ if the defect is in the metal. These two are represented in Fig.3.2(a). We have reported that for $x >> \lambda$ and $Z_+ \sim 0$, $\hat{\mathbf{G}}_I^{(i,i)}$ tends to the value of the residue alone $\hat{\mathbf{G}}_p^{(i,i)}$. In general the plasmon modes also have a skin-depth penetration in the metal. This can be calculated through the transmitted Green Tensor that propagate the field from air to metal $\hat{\mathbf{G}}_p^{(1,2)}$. Viceversa $\hat{\mathbf{G}}_p^{(2,1)}$ would give the penetration of the surface polariton into the air semi-space. These terms are the residues of the transmitted Green tensors $\hat{\mathbf{G}}_I^{(1,2)}$ and $\hat{\mathbf{G}}_I^{(2,1)}$ represented in Fig.3.2(b). The transmitted Green tensor has been omitted so far because it has no relevance to the the calculations of the fields in the source region of the inhomogeneity, nonetheless it is useful in the calculation of the fields far from the sources. The two Green tensors $\hat{\mathbf{G}}^{(1,2)}$ and $\hat{\mathbf{G}}^{(2,1)}$ can be simply attained by the method of scattering superposition[2],

by matching the same electric field boundary conditions as $\hat{\mathbf{G}}^{(1,1)}$ and $\hat{\mathbf{G}}^{(2,2)}$, respectively, but on the opposite side of the interface (look up Sec.1.13). For instance if we have a source in air semi-space(1), we can calculate the field emitted in the metal with $\hat{\mathbf{G}}^{(1,2)}$:

$$\hat{\mathbf{G}}^{(1,2)}(\boldsymbol{\rho},\boldsymbol{\rho}') = \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty}\frac{dk_x}{k_z}\, e^{ik_x(x-x')}e^{i(k_z z - k_z^m z')}\left[t_p^{(1,2)}\,\mathbf{k}_p^{-}\mathbf{k}_p^{m-} + t_s^{(1,2)}\mathbf{k}_s\mathbf{k}_s\right] \quad (3.5)$$

Define $\hat{\mathbf{\Gamma}}_p^{(1,2)}(k_x) = \mathbf{k}_p^{-}\mathbf{k}_p^{m-}$:

$$\hat{\mathbf{\Gamma}}_p^{(1,2)}(k_x) = \frac{k_z k_z^m}{g^2\,\sqrt{\varepsilon}k_\parallel^2}\begin{pmatrix} {k_x}^2 & k_x k_y & k_x {k_\parallel}^2/k_z^m \\ k_x k_y & k_y^2 & {k_\parallel}^2 k_y/k_z^m \\ k_x {k_\parallel}^2/k_z & {k_\parallel}^2 k_y/k_z & {k_\parallel}^4/(k_z^m\, k_z) \end{pmatrix} \quad (3.6)$$

Define $\hat{\mathbf{\Gamma}}_s^{(1,2)}(k_x) = \mathbf{k}_s\mathbf{k}_s$:

$$\hat{\mathbf{\Gamma}}_s(k_x) = \frac{1}{k_\rho^2}\begin{pmatrix} k_y^2 & -k_x k_y & 0 \\ -k_x k_y & k_x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.7)$$

Likewise for a source in the metal its propagation to the air region is given by:

$$\hat{\mathbf{G}}^{(2,1)}(\boldsymbol{\rho},\boldsymbol{\rho}') = \frac{i}{4\pi}\int_{-\infty}^{\infty}\frac{dk_x}{k_z^m}\, e^{ik_x(x-x')}e^{i(k_z z - k_z^m z')}\left[t_p^{(2,1)}\,\mathbf{k}_p^{m+}\mathbf{k}_p^{+} + t_s^{(2,1)}\mathbf{k}_s\mathbf{k}_s\right] \quad (3.8)$$

now define $\hat{\mathbf{\Gamma}}_p^{(2,1)}(k_x) = \mathbf{k}_p^{m+}\mathbf{k}_p^{+}$:

$$\hat{\mathbf{\Gamma}}_p^{(2,1)}(k_x) = \frac{k_z k_z^m}{g^2\,\sqrt{\varepsilon}k_\parallel^2}\begin{pmatrix} {k_x}^2 & k_x k_y & -k_x {k_\parallel}^2/k_z^m \\ k_x k_y & k_y^2 & -{k_\parallel}^2 k_y/k_z^m \\ -k_x {k_\parallel}^2/k_z & -{k_\parallel}^2 k_y/k_z & {k_\parallel}^4/(k_z^m\, k_z) \end{pmatrix} \quad (3.9)$$

Having defined the transmitted Green Tensors we are now able to reproduce analytical expressions for the plasmon excitation, with applying the residue theorem as we did in the last chapter. Deriving the residue contribution that comes from the surface plasmon resonance of $\hat{\mathbf{G}}^{(1,2)}$ (when the source in the air semi-space) is trivial, if we first look up the case of $\hat{\mathbf{G}}^{(1,1)}$ in Eq.(2.22). We know that the final result of the residue theorem gives:

$$\hat{\mathbf{G}}_p^{(1,1)} = G\, e^{ik_{px}\,x_-}\, e^{ik_{pz}\,(z+z')}\;\hat{\mathbf{\Gamma}}_{spp}^{(1,1)} \quad (3.10)$$

where is given $\hat{\mathbf{\Gamma}}^{(1,1)}_{spp}$ is given in Eq.(2.26). Similarly applying the residue theorem to Eq.(3.5) we end up with:

$$\hat{\mathbf{G}}^{(1,2)}_{p} = G\, e^{ik_{px}\, x_-}\; e^{i(k_{pz}\, z + k^{m}_{zp}\, z')}\;\; \hat{\mathbf{\Gamma}}^{(1,2)}_{spp} \tag{3.11}$$

where $\hat{\mathbf{\Gamma}}^{(1,2)}_{spp} = \hat{\mathbf{\Gamma}}^{(1,2)}_{p}(k_{px})$ and the pre-factor G is given in Eq.(2.23):

$$G = \frac{-\varepsilon^2}{\varepsilon^2 - 1} \frac{k^3_{pz}}{k^2_p\, g^2 k_{px}}. \tag{3.12}$$

Recall that $k_p = g\sqrt{\frac{\varepsilon}{\varepsilon+1}}$ and $k_{px} = \sqrt{k^2_p - k^2_y}$.
Notice G is a common constant, to all plasmon residues so far. As we are going to prove in Sec.3.3.1, this is not coincidental; in fact G is related to the energy of the surface plasmon in a lossless metal.

3.3 The Surface Plasmon Field in the 2D Green Tensor

Now we are going highlight a rearrangement of $\hat{\mathbf{G}}_p^{(1,1)}$ and $\hat{\mathbf{G}}_p^{(1,2)}$, that will allow us to write the excitation of surface plasmon caused by a point source in an air-metal back-ground. At the metal interface $z = 0$ all fields are discontinuous; it is convenient, however, to generalize the field of a surface plasmon polariton in a form valid at both sides of the interface. This indeed emerges naturally out of the mathematics of the Green tensor formalism as a way of generalizing the pole in the Green tensor in a compact form.
Let us start from the case of a source above the surface. Let us define θ such that $k_y = k_p \sin\theta$ and hence $k_{px} = k_p \cos\theta$, as in Fig.3.1 and Fig.3.3. Also define the following quantity:

$$Z_s = \frac{k_{pz}}{k_p} = \frac{1}{\sqrt{\varepsilon}} \tag{3.13}$$

which is known as surface impedance. To condense the notation take a point

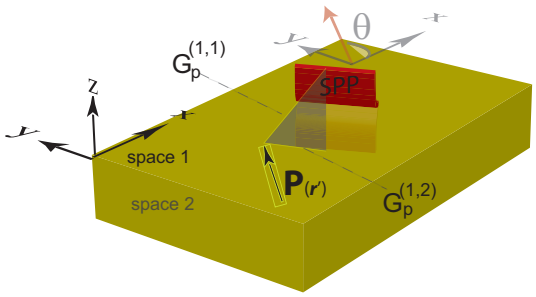

Figure 3.3: A schematic representation of the Green tensor residues $\hat{\mathbf{G}}_p^{(1,1)}$ and $\hat{\mathbf{G}}_p^{(1,2)}$. $\hat{\mathbf{G}}_p^{(1,1)}$ quantifies the emission from a point-polarization element to the electric field of a SPP in the air semi-space. $\hat{\mathbf{G}}_p^{(1,2)}$ quantifies the emission from a point-polarization element to the electric field of a SPP in the metal semi-space.

polarization source element $\mathbf{P}(\boldsymbol{\rho}')$, located at $\boldsymbol{\rho}'$ with $z' > 0$ (above the surface), which is an arbitrary vector in the air space see Fig.3.3. Its net emission into surface plasmons is attained from the Green tensors penetration into both air and metal. The Green tensors that give such penetration are $\hat{\mathbf{G}}_p^{(1,1)}$ and $\hat{\mathbf{G}}_p^{(1,2)}$. However from Eq.(2.24) we have that:

$$\hat{\mathbf{G}}_p^{(1,1)} \cdot \mathbf{P}(\boldsymbol{\rho}') \propto \begin{pmatrix} Z_s \cos\theta \\ Z_s \sin\theta \\ 1 \end{pmatrix}, \tag{3.14}$$

from Eq.(3.11) we have that:

$$\hat{\mathbf{G}}_p^{(1,2)} \cdot \mathbf{P}(\boldsymbol{\rho}') \propto \begin{pmatrix} Z_s \cos\theta \\ Z_s \sin\theta \\ Z_s^2 \end{pmatrix}. \tag{3.15}$$

Their emission results in a vector $\hat{\mathbf{G}}_p \cdot \mathbf{P}$ that can be now written in a compact form valid for the whole of space. To do so, we just need to define the electric field and Poynting vector of a surface plasmon.

3.3.1 Surface Plasmon Polariton Mode

From Eq.(3.14) and Eq.(3.15) we can create a generalized vector $\hat{\mathbf{G}}_p \cdot \mathbf{P}$, if we define the field: $\mathbf{e}_{spp\pm}(\mathbf{r}) = \mathbf{e}_{spp\pm}^{(\nu=1)}$ for $z > 0$, and $\mathbf{e}_{spp\pm}(\mathbf{r}) = \mathbf{e}_{spp\pm}^{(\nu=2)}$ for $z < 0$; where:

$$\mathbf{e}_{spp\pm}^{(\nu=1)}(\mathbf{r}) = \begin{pmatrix} \pm Z_s \cos\theta \\ Z_s \sin\theta \\ 1 \end{pmatrix} e^{i(\pm k_{px} x + k_y y + k_{pz} z)}, \quad z > 0, \tag{3.16}$$

$$\mathbf{e}_{spp\pm}^{(\nu=2)}(\mathbf{r}) = \begin{pmatrix} \pm Z_s \cos\theta \\ Z_s \sin\theta \\ Z_s^2 \end{pmatrix} e^{i(\pm k_{px} x + k_y y - k_{pz}^m z)}, \quad z < 0. \tag{3.17}$$

This mode correspond to a monochromatic surface plasmon propagating in the positive x direction ($\mathbf{e}_{spp+}(\mathbf{r})$) or negative ($\mathbf{e}_{spp-}(\mathbf{r})$). Furthermore $\mathbf{e}_{spp+}(\mathbf{r})$ is connected to the constant G of Eq.(3.12) which is common to the pole of the Green tensor on any side of the interface.
Let us calculate the incident Poynting energy flux of the surface plasmon mode. The magnetic field of the surface plasmon mode is continuous at the interface and equal to:

$$\mathbf{h}_{spp\pm}(\mathbf{r}) = -\frac{\mathtt{i}}{g} \nabla \times \mathbf{e}_{spp\pm}(\mathbf{r}). \tag{3.18}$$

Now consider a lossless metal with ε real and negative. The Poynting vector is:

$$\mathbf{e}_{spp+} \times \mathbf{h}_{spp+}^* \cdot \mathbf{u}_x = e^{-2|k_{pz}|z} \, \frac{k_p}{g} \left[Z_s^2 + 1\right] \cos\theta \qquad z > 0 \tag{3.19}$$

$$\mathbf{e}_{spp+} \times \mathbf{h}_{spp+}^* \cdot \mathbf{u}_x = Z_s^2 \, e^{-2|k_{pz}^m|z} \, \frac{k_p}{g} \left[Z_s^2 + 1\right] \cos\theta \qquad z < 0 \tag{3.20}$$

so that the energy fluxes of the surface plasmon modes through a plane wavefront, can be calculated from the fluxes in air and metal respectively:

$$\begin{aligned} S_{spp}^{(\nu=1)}(\theta) &= \int_0^\infty dz\, \mathbf{e}_{spp+} \times \mathbf{h}_{spp+}^* \cdot \mathbf{u}_x = \frac{k_p}{g} \frac{Z_s^2+1}{2|k_{pz}|} \cos\theta \\ S_{spp}^{(\nu=2)}(\theta) &= \int_{-\infty}^0 dz\, \mathbf{e}_{spp} \times \mathbf{h}_{spp}^* \cdot \mathbf{u}_x = Z_s^4\, S_{spp}^{(\nu=1)}(\theta). \end{aligned}$$

Finally the total Poynting vector energy flux associated to a plasmon mode in a lossless metal is:

$$\begin{aligned} S_{spp}(\theta) &= \int_{-\infty}^\infty dz'\, Re\{\mathbf{e}_{spp} \times \mathbf{h}_{spp}^* \cdot \mathbf{u}_x\} = -\frac{\mathtt{i}}{2g} \left(\frac{k_{pz}^2}{G k_p^4} \right) = \\ &= \frac{\sqrt{-\varepsilon}}{2g} \frac{(\varepsilon+1)(\varepsilon^2-1)}{\varepsilon^3} \cos\theta > 0 \end{aligned}$$

Finally, at normal incidence when $\theta = 0$, let us define: $S_{spp} = \frac{\sqrt{-\varepsilon}}{2g} \frac{(\varepsilon+1)(\varepsilon^2-1)}{\varepsilon^3}$. These parameters are calculated for a free-running surface plasmon mode in the unperturbed background. The form of $\mathbf{e}_{spp+}(\mathbf{r})$ is however relevant to the surface plasmon excited in 3D problems, where k_y and θ are unconstrained parameters. However we are mostly interested in 2D systems, where as explained in Sec.3.1, k_y and θ are conserved and fixed in all fields. That includes the excited surface plasmon modes. Correspondingly the y-dependence, will be implicit in 2D systems where the surface plasmon field is expressed as $\mathbf{e}_{spp\pm}(\boldsymbol{\rho})$.

3.4 Emission of Surface Plasmon Polaritons

3.4.1 2D Source above the Metal

Having defined a surface plasmon mode we can condense the emission of a point polarization element $\mathbf{P}(\boldsymbol{\rho}')$ above the metal, into both air $\hat{\mathbf{G}}_p^{(1,1)}$ and metal $\hat{\mathbf{G}}_p^{(1,2)}$, through the field of a surface plasmon, as:

$$\hat{\mathbf{G}}_{p+}(\boldsymbol{\rho},\boldsymbol{\rho}')\cdot\mathbf{P}(\boldsymbol{\rho}') = -\tilde{Z}\frac{k_p}{k_{px}}\, e^{i(k_{pz}z'-k_{px}x')}\, \left[\left(-Z_s\mathbf{u}_{\parallel}+\mathbf{u}_z\right)\cdot\mathbf{P}(\boldsymbol{\rho}')\right]\, \mathbf{e}_{spp+}(\boldsymbol{\rho}),$$

with $\tilde{Z} = Z_s((1+Z_s^2)(Z_s^4-1))^{-1}$.
Notice these expressions are valid for $k_y \neq 0$. The dependence on $k_y = k_p \sin\theta$ is contained in the factors, k_p/k_{px}, $\mathbf{u}_{\parallel} = \cos\theta\mathbf{u}_x + \sin\theta\mathbf{u}_y$ and $\mathbf{e}_{spp+}^{(\nu=1)}(\boldsymbol{\rho})$ (see Eq.(3.16)).
If the back-ground metal is lossless ($\varepsilon \in \Re$):

$$\tilde{Z}\frac{k_p}{k_{px}} = \frac{\mathtt{i}}{2g\, S_{spp}(\theta)} \tag{3.21}$$

and thus:

$$\hat{\mathbf{G}}_p(\boldsymbol{\rho},\boldsymbol{\rho}')\cdot\mathbf{P}(\boldsymbol{\rho}') \;=\; -\frac{\mathtt{i}}{2g\, S_{spp}(\theta)}\; \left[\left(\mathbf{e}_{spp}^{(\nu=1)}(\boldsymbol{\rho}')\right)^{*}\cdot\mathbf{P}(\boldsymbol{\rho}')\right]\, \mathbf{e}_{spp}(\boldsymbol{\rho}'). \tag{3.22}$$

Correspondingly the field emitted into surface plasmon modes, by source distribution $\mathbf{P}(\boldsymbol{\rho}')$ extended over an area A (above the metal) is:

$$\mathbf{E}_{p+}(\boldsymbol{\rho}) \;=\; E_p\;\mathbf{e}_{spp+}(\boldsymbol{\rho}) \tag{3.23}$$

$$E_{p+} \;=\; -g^2\tilde{Z}\frac{k_p}{k_{px}}\int_A d\boldsymbol{\rho}'\, e^{i(k_{pz}z'-k_{px}x')}\, \left(-Z_s\mathbf{u}_{\parallel}+\mathbf{u}_z\right)\cdot\mathbf{P}(\boldsymbol{\rho}'). \tag{3.24}$$

while for the lossless case:

$$E_{p+} \;=\; -\frac{\mathtt{i}g}{2\, S_{spp}(\theta)}\int_A d\boldsymbol{\rho}'\, \mathbf{e}_{spp+}^{*}(\boldsymbol{\rho}')\cdot\mathbf{P}(\boldsymbol{\rho}') \tag{3.25}$$

Furthermore since:

$$\mathbf{H}_{p+}(\boldsymbol{\rho}) = E_{p+}\,\mathbf{h}_{spp+}(\boldsymbol{\rho}), \tag{3.26}$$

if the the source $\mathbf{P}(\boldsymbol{\rho}')$ is produced by an incident surface plasmon field, we can define the scattering energy flux normalized to the incident energy as:

$$\sigma_p^{+} = \frac{\mathbf{E}_{p+}\times\mathbf{H}_{p+}^{*}\cdot\mathbf{u}_x}{\mathbf{e}_{spp+}\times\mathbf{h}_{spp+}^{*}\cdot\mathbf{u}_x} = |E_{p+}|^2. \tag{3.27}$$

Everything we just said is also valid mutantis mutandis for the scattering into a backwards plasmon. The related scattering cross-section is:

$$\sigma_p^- = \frac{\mathbf{E}_{p-} \times \mathbf{H}^*_{p-} \cdot \mathbf{u}_x}{\mathbf{e}_{spp-} \times \mathbf{h}^*_{spp-} \cdot \mathbf{u}_x} = |E_{p-}|^2 \tag{3.28}$$

where:

$$E_{p-} = -g^2 \tilde{Z} \frac{k_p}{k_{px}} \int_A d\boldsymbol{\rho}' \; e^{i(k_{pz} z' + k_{px} x')} \left[Z_s P_x(\boldsymbol{\rho}') \cos\theta - Z_s P_y(\boldsymbol{\rho}') \sin\theta + P_z(\boldsymbol{\rho}') \right],$$

and this simplifies once again, for the lossless case, to:

$$E_{p-} \;=\; -\frac{\mathtt{i}g}{2\, S_{spp}(\theta)} \int_A d\boldsymbol{\rho}' \; \mathbf{e}^*_{spp-}(\boldsymbol{\rho}') \cdot \mathbf{P}(\boldsymbol{\rho}'). \tag{3.29}$$

3.4.2 Source below the Metal

Likewise, for a source below the surface, repeating the same procedure we end up with:

$$\sigma_p^{\pm} = \frac{\mathbf{E}_{p\pm} \times \mathbf{H}^*_{p\pm} \cdot \mathbf{u}_x}{\mathbf{e}_{spp\pm} \times \mathbf{h}^*_{spp\pm} \cdot \mathbf{u}_x} = |E_{p\pm}|^2 \tag{3.30}$$

since:

$$\mathbf{E}_{p\pm}(\boldsymbol{\rho}) \;=\; E_{p\pm} \; \mathbf{e}_{spp\pm}(\boldsymbol{\rho}) \tag{3.31}$$

$$E_{p+} \;=\; -g^2 \tilde{Z} \frac{k_p}{k_{px}} \int_A d\boldsymbol{\rho}' \; e^{i(k^m_{pz} z' - k_{px} x')} \; \left(-Z_s \mathbf{u}_{\parallel} + Z_s^2 \mathbf{u}_z \right) \cdot \mathbf{P}(\boldsymbol{\rho}')$$

$$E_{p-} = -g^2 \tilde{Z} \frac{k_p}{k_{px}} \int_A d\boldsymbol{\rho}' \; e^{i(k^m_{pz} z' + k_{px} x')} \left[Z_s P_x(\boldsymbol{\rho}) \cos\theta - Z_s P_y(\boldsymbol{\rho}) \sin\theta + Z_s^2 P_z(\boldsymbol{\rho}) \right]$$

which for the lossless case:

$$E_{p\pm} \;=\; -\frac{\mathtt{i}g}{2\, S_{spp}(\theta)} \int_A d\boldsymbol{\rho}' \; \mathbf{e}^*_{spp\pm}(\boldsymbol{\rho}') \cdot \mathbf{P}(\boldsymbol{\rho}'). \tag{3.32}$$

3.4.3 3D source Analog

In this section we point out, briefly, the analog result for the surface plasmon polariton scattering from a source in a 3D system. We will need this particular result in the Chapter 5, to explain the existence of a Brewster-type angle for the scattering of surface plasmon polaritons.
The coupling of a dipole with a 3D system is based on the contribution of the surface plasmon pole to the net three-dimensional Green tensor. Its derivation in 3D is extremely similar to the one we have just carried out in 2D. Thus we shall not repeat it, but it can be found for reference in [36]. Let us consider a point polarization element at the origin of our reference system $\mathbf{P}(\mathbf{r}' = 0) = \mathbf{P}$. The expressions of the surface plasmon coupling to of such source in a 3D system are:
a) For a dipole placed in the vacuum side of the metal-vacuum interface:

$$\mathbf{E}_p(\mathbf{r}) = \tilde{Z} g^2 \sqrt{\frac{k_p}{2\pi \mathtt{i} R_{\|}}} \left[Z_s \, \mathbf{P} \cdot \mathbf{u}_{R_{\|}} - \mathbf{P} \cdot \mathbf{u}_z \right] \mathbf{e}_{spp}(\mathbf{r}), \tag{3.33}$$

b) for a dipole placed in the metal side of the metal-vacuum interface:

$$\mathbf{E}_p(\mathbf{r}) = \tilde{Z} g^2 \sqrt{\frac{k_p}{2\pi \mathtt{i} R_{\|}}} \left[Z_s \, \mathbf{P} \cdot \mathbf{u}_{R_{\|}} - Z_s^2 \, \mathbf{P} \cdot \mathbf{u}_z \right] \mathbf{e}_{spp}(\mathbf{r}), \tag{3.34}$$

where $\mathbf{u}_{R_{\|}} = \cos\phi \mathbf{u}_x + \sin\phi \mathbf{u}_y$, and where $k_{px}\mathbf{u}_x + k_y \mathbf{u}_y = k_p \mathbf{u}_{R_{\|}}$. Therefore, a dipole oriented along the interface emits plasmons preferentially in the longitudinal direction, and does not radiate SPPs in the perpendicular direction. By contrast a dipole placed at the surface, but oriented along its normal direction, radiates plasmons isotropically in the plane.

3.5 The Far-Field Scattered by 2D Systems

In this section we shall obtain the rest of the contribution to the scattered field, coming from all of the other modes. This takes on a particularly simple form far from the source where: $\boldsymbol{\rho} >> \boldsymbol{\rho}'$ and $\boldsymbol{\rho} >> \lambda$ (also $x >> x'$ and $z >> z'$). Here the Method of the Steepest Descent can be applied to get an analytic expression for the Green Tensor. In general all of our Sommerfeld integrals representing $\hat{\mathbf{G}}^{(i,j)}$, can be written as:

$$\hat{\mathbf{G}}(\boldsymbol{\rho}, \boldsymbol{\rho}') = \int_{-\infty}^{\infty} dk_x \, e^{iF(k_x, x-x', z, z')} \, \hat{\mathbf{\Gamma}}(k_x) \tag{3.35}$$

where

$$F(k_x, x - x', z, z') = k_x(x - x') + k_z z + k_z^{(\nu)} z' \tag{3.36}$$

and where in the cases we are going to treat: $k_z^{(1)} = k_z$ and $k_z^{(2)} = k_z^m$. Now according to the prescription of the method of the steepest descent[1], as ρ approaches infinity, the exponential term oscillates very fast and cancels the contribution from the rest of the integral, Eq.(3.35), at every cycle. The only

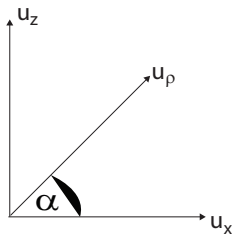

Figure 3.4: A schematic representation of the angle α, and of the basis vectors $\mathbf{u}_x$, $\mathbf{u}_z$, $\mathbf{u}_\rho$.

contribution, at an infinite distance from the source, is thus coming from the spectral zone in the neighborhood of which the argument of the exponential $F(k_x, x - x', z, z')$ has an extremum, as the exponential is the least oscillatory in such neighborhood. We can determine this contribution by setting $F(k_x, x - x', z, z')$ at its critical point:

$$\frac{\partial}{\partial k_x} F(k_x, x - x', z, z') = 0. \tag{3.37}$$

Remarkably the solution corresponds the limit of the diffraction problem (of Eq.(3.35)), in which geometrical optics is recovered. *In the far-field light propagates on a straight line.* If we express the point $\boldsymbol{\rho}$ in polar coordinates (ρ, α) as:

$$\boldsymbol{\rho} = \rho \mathbf{u}_\rho = \rho \cos\alpha \, \mathbf{u}_x + \rho \sin\alpha \, \mathbf{u}_z \tag{3.38}$$

we can write the solution to Eq.(3.37) as:

$$\boldsymbol{\kappa} = g\,\mathbf{u}_\rho \tag{3.39}$$

where $\boldsymbol{\kappa} = k_x^0\mathbf{u}_x + k_z^0\mathbf{u}_z$ and correspondingly we have designated the value of the variables at critical points as: $k_x^0 = \kappa\cos\alpha$ and $k_z^0 = \kappa\sin\alpha$ (remember that by definition $\kappa = \sqrt{g^2 - k_y^2}$ see Sec.1.10).
Next the Green Tensor is reduced to a function of the minimal optical path by expanding $F(k_x, x-x', z, z')$ in a second term taylor series. For the sake of tidiness we shall write $F(k_x, x-x', z, z')$ simply as $F(k_x)$ but it is meant that the dependence of the rest of the variables is the one given in Eq.(3.36)

$$\begin{aligned}\hat{\mathbf{G}}(\boldsymbol{\rho},\boldsymbol{\rho}') &= e^{iF(k_x^0)}\,\hat{\mathbf{\Gamma}}(k_x^0)\int_{-\infty}^{\infty} dk_x\,\exp\left[-\frac{1}{2}\mathtt{i}F''(k_x^0)(k_x-k_x^0)^2\right] = \\ &= e^{i\pi/4}e^{iF(k_x^0)}\,\hat{\mathbf{\Gamma}}(k_x^0)\sqrt{\frac{2\pi}{F''(k_x^0)}}\end{aligned} \tag{3.40}$$

where $F''(k_x^0) = \left[\frac{\partial^2}{\partial k_x^2}F(k_x)\right]_{k_x=k_x^0}$ and where we have used the identity: $\int_{-\infty}^{\infty} d\zeta e^{\mathtt{i}\zeta^2} = e^{i\pi/4}\sqrt{\pi}$.
Consider now the following Green tensors. The free-space 2D Green tensor $\hat{\mathbf{G}}_0$ defined as the homogeneous medium Green tensor of Sec.1.10, with $\varepsilon_b = 1$. The reflected 2D Green tensor $\hat{\mathbf{G}}_I^{(1,1)}$ defined in Eq.(1.112). The refracted 2D Green tensor $\hat{\mathbf{G}}^{(2,1)}$ defined in Eq.(3.8). The three tensors are represented schematically in Fig.3.5. Applying the method of the steepest descent to the total Green tensors $\hat{\mathbf{G}}^{(1,1)} = \hat{\mathbf{G}}_0 + \hat{\mathbf{G}}_I^{(1,1)}$ and $\hat{\mathbf{G}}^{(2,1)}$, we get:

3.5.1 Direct GT

In free-space:

$$\begin{aligned}\hat{\mathbf{G}}_0(\boldsymbol{\rho}-\boldsymbol{\rho}') &= \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty}\frac{dk_x}{k_z}\,e^{ik_z(z-z')}e^{ik_x(x-x')}\,\left(\mathbf{k}_p^+\mathbf{k}_p^+ + \mathbf{k}_s\mathbf{k}_s\right) \\ \lim_{\rho\to\infty}\hat{\mathbf{G}}_0(\boldsymbol{\rho}-\boldsymbol{\rho}') &= e^{i\pi/4}e^{-i\kappa(x'\cos\alpha+z'\sin\alpha)}\,\frac{e^{i\kappa\rho}}{\sqrt{8\pi\kappa\,\rho}}\,\left[\mathbf{k}_p^+\mathbf{k}_p^+ + \mathbf{k}_s\mathbf{k}_s\right]_{k_x=k_x^0}\end{aligned}$$

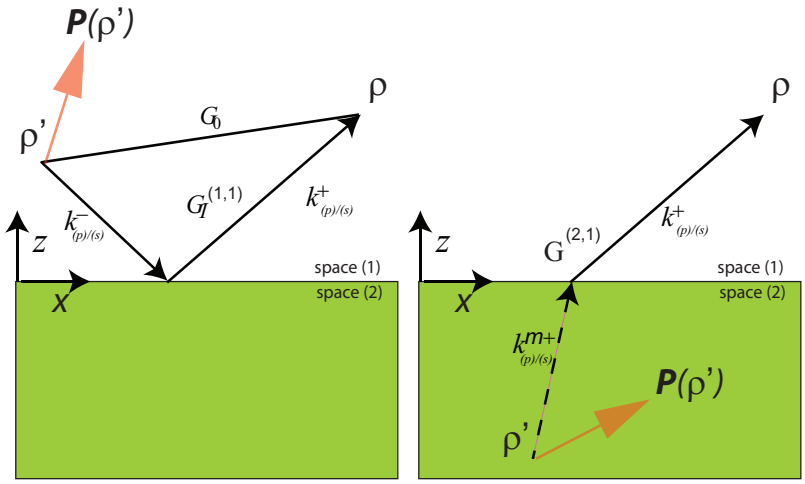

Figure 3.5: The radiation from a point-polarization element above a metal slab is obtained through the direct Green tensor $\hat{\mathbf{G}}_0$ and the indirect Green tensor $\hat{\mathbf{G}}_I^{(1,1)}$. The radiation from a point-polarization element inside a metal slab is obtained through the transmitted Green tensor $\hat{\mathbf{G}}^{(2,1)}$.

3.5.2 Reflected GT source upon the metal

$$\hat{\mathbf{G}}_I^{(1,1)}(\boldsymbol{\rho},\boldsymbol{\rho}') = \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{e^{ik_x(x-x')}e^{ik_z(z+z')}}{k_z}\left(r_p(k_x)\,\mathbf{k}_p^-\mathbf{k}_p^+ + r_s(k_x)\mathbf{k}_s\mathbf{k}_s\right)$$

$$\lim_{\rho\to\infty}\hat{\mathbf{G}}_I^{(1,1)}(\boldsymbol{\rho},\boldsymbol{\rho}') = e^{i\pi/4}e^{-i\kappa(x'\cos\alpha - z'\sin\alpha)}\frac{e^{i\kappa\rho}}{\sqrt{8\pi\kappa\,\rho}}\left[r_p\,\mathbf{k}_p^-\mathbf{k}_p^+ + r_s\mathbf{k}_s\mathbf{k}_s\right]_{k_x=k_x^0}$$

3.5.3 Refracted GT (Source in the metal)

$$\hat{\mathbf{G}}^{(2,1)}(\boldsymbol{\rho},\boldsymbol{\rho}') = \frac{\mathtt{i}}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{e^{i(k_z z - k_z^m z')}e^{ik_x(x-x')}}{k_z^m}\left(t_p^{(2,1)}(k_x)\mathbf{k}_p^{m+}\mathbf{k}_p^+ + t_s^{(2,1)}(k_x)\mathbf{k}_s\mathbf{k}_s\right)$$

$$\begin{aligned}\lim_{\rho\to\infty}\hat{\mathbf{G}}^{(2,1)}(\boldsymbol{\rho},\boldsymbol{\rho}') &= e^{i\pi/4}e^{-i\kappa\,x'\cos\alpha}e^{-ik_z^m z'}\times\\ &\times\ \frac{e^{i\kappa\rho}}{\sqrt{8\pi\kappa\,\rho}}\frac{k_z^m}{k_z}\left[t_p^{(2,1)}\mathbf{k}_p^{m+}\mathbf{k}_p^+ + t_s^{(2,1)}\mathbf{k}_s\mathbf{k}_s\right]_{k_x=k_x^0} =\\ &= e^{i\pi/4}e^{-i\kappa\,x'\cos\alpha}e^{-ik_z^m z'}\times\\ &\times\ \frac{e^{i\kappa\rho}}{\sqrt{8\pi\kappa\,\rho}}\left[t_p^{(1,2)}\,\mathbf{k}_p^{m+}\mathbf{k}_p^+ + t_s^{(1,2)}\mathbf{k}_s\mathbf{k}_s\right]_{k_x=k_x^0}.\end{aligned}$$

3.6 Far Fields at oblique incidence

We are now in a position to write the far-fields propagated by a source distribution $\mathbf{P}(\boldsymbol{\rho}')$ extended over an area A.
At oblique incidence, $k_y \neq 0$, in air p-modes are:

$$\mathbf{k}_p^{\pm} = \frac{\sin\alpha}{k_{\parallel}} \left[g\cos\alpha \mathbf{u}_x + k_y \mathbf{u}_y \mp \frac{k_{\parallel}^2}{g\sin\alpha} \mathbf{u}_z \right] \tag{3.41}$$

with $k_{\parallel} = \sqrt{g\cos\alpha^2 + k_y^2}$.

In the metal:

$$\mathbf{k}_p^{m\pm} = \frac{k_z^m}{g\sqrt{\varepsilon} k_{\parallel}} \left[g\cos\alpha \mathbf{u}_x + k_y \mathbf{u}_y \mp \frac{k_{\parallel}^2}{k_z^m} \mathbf{u}_z \right]. \tag{3.42}$$

and where $k_z^m = \sqrt{\varepsilon g^2 - k_y^2 - g^2\cos^2\alpha}$.
Finally s-modes are:

$$\mathbf{k}_s = \frac{1}{k_{\parallel}} \left[\, g\cos\alpha \mathbf{u}_x - k_y \mathbf{u}_y \, \right] \tag{3.43}$$

Now going back to Eq.(3.4), the net field in the far zone can be written as:

$$\mathbf{E}(\boldsymbol{\rho}) = \mathbf{E}_b(\boldsymbol{\rho}) + \mathbf{E}_{rad}(\boldsymbol{\rho}). \tag{3.44}$$

Using the far-field Green tensor derived in the last section we have an analytical expression for the radiative field in the far zone $\mathbf{E}_{rad}(\boldsymbol{\rho})$:

$$\mathbf{E}_{rad}(\boldsymbol{\rho}) = g^2 \frac{e^{i(\kappa\rho - \pi/4)}}{\sqrt{8\pi\kappa\rho}} \left[E_{rad}^{(p)}(\alpha)\, \mathbf{k}_p^{+} + E_{rad}^{(s)}(\alpha)\, \mathbf{k}_s \right] \tag{3.45}$$

For a source in the air and using $\hat{\mathbf{G}}_b = \hat{\mathbf{G}}_0 + \hat{\mathbf{G}}_I^{(1,1)}$ in Eq.(3.4) we get Eq.(3.45) with:

$$E_{rad}^{(p)}(\alpha) = \int_A d\boldsymbol{\rho}' e^{-i\kappa \mathbf{u}_\rho \cdot \boldsymbol{\rho}'} \left[(\mathbf{k}_p^{+} + \mathbf{k}_p^{-}\, r_p\, e^{2i\kappa z' \sin\alpha}) \cdot \mathbf{P}(\boldsymbol{\rho}') \right] \tag{3.46}$$

$$E_{rad}^{(s)}(\alpha) = \int_A d\boldsymbol{\rho}' e^{-i\kappa \mathbf{u}_\rho \cdot \boldsymbol{\rho}'} \left[1 + r_s\, e^{2i\kappa z' \sin\alpha} \right] (\mathbf{k}_s \cdot \mathbf{P}(\boldsymbol{\rho}')). \tag{3.47}$$

Likewise for a source in the metal, using $\hat{\mathbf{G}}_b = \hat{\mathbf{G}}^{(2,1)}$ in Eq.(3.4) we get Eq.(3.45) with:

$$E_{rad}^{(p)}(\alpha) = \int_A d\boldsymbol{\rho}' e^{-i\kappa\, x' \cos\alpha} e^{-ik_z^m z'}\, t_p^{(1,2)} \left(\mathbf{k}_p^{m+} \cdot \mathbf{P}(\boldsymbol{\rho}') \right) \tag{3.48}$$

$$E_{rad}^{(s)}(\alpha) = \int_A d\boldsymbol{\rho}' e^{-i\kappa\, x' \cos\alpha} e^{-ik_z^m z'}\, t_s^{(1,2)} \left(\mathbf{k}_s \cdot \mathbf{P}(\boldsymbol{\rho}') \right) \tag{3.49}$$

Despite their tedious expressions as an function of α, the modes $\mathbf{k}_p^{\pm}$,$\mathbf{k}_p^{m\pm}$ and $\mathbf{k}_s$ preserve the important property that the electric field far from the source is always transversal to its direction of propagation. Consider in fact the far-field wave-vector :

$$\mathbf{k} = k_x\mathbf{u}_x + k_y\mathbf{u}_y + k_z\mathbf{u}_z = \kappa\cos\alpha\mathbf{u}_x + k_y\mathbf{u}_y + \kappa\sin\alpha\mathbf{u}_z \tag{3.50}$$

it is easy to verify that:

$$\mathbf{k}\cdot\mathbf{k}_p^{\pm} = \mathbf{k}\cdot\mathbf{k}_p^{m\pm} = \mathbf{k}\cdot\mathbf{k}_s = 0 \tag{3.51}$$

in fact the three wavevectors $\mathbf{k}/g, \mathbf{k}_p^b, \mathbf{k}_s$ constitute an orthogonal basis, as explained in Sec.1.2. It is important to remark at this stage that thanks to the property Eq.(3.51):

$$\mathbf{E}_{rad}(\rho,\alpha)\cdot\mathbf{k} = 0 \tag{3.52}$$

3.7 Radiative Energy at Oblique Incidence

Using the property of Eq.(3.52), we can greatly simplify the calculation of the Poynting vector in the far-field, which would otherwise have a formidable expression at oblique incidence.
Consider the cylindrical right-handed orthogonal basis:

$$\mathbf{u}_y, \quad \mathbf{u}_\rho = \cos\alpha \mathbf{u}_x + \sin\alpha \mathbf{u}_z, \quad \mathbf{u}_\alpha = -\sin\alpha \mathbf{u}_x + \cos\alpha \mathbf{u}_z \tag{3.53}$$

and the wavevectors:

$$\mathbf{k} = \kappa\, \mathbf{u}_\rho + k_y \mathbf{u}_y \quad \mathbf{u}_k = \mathbf{k}/g \tag{3.54}$$

From Maxwell's equations the magnetic field is:

$$\mathbf{H} = -\frac{i}{g} \nabla \times \mathbf{E} = \mathbf{u}_k \times \mathbf{E} \tag{3.55}$$

and the related Poynting vector is:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}^* = \mathbf{E} \times (\mathbf{u}_k \times \mathbf{E}^*) \tag{3.56}$$

Now we can find a shortcut to computing $\mathbf{S}$, if we notice the following identity:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}. \tag{3.57}$$

Hence the Poynting vector reads:

$$\mathbf{S} = |\mathbf{E}|^2 \mathbf{u}_k - (\mathbf{E} \cdot \mathbf{u}_k)\mathbf{E}^* \tag{3.58}$$

However, since $\mathbf{k} \cdot \mathbf{E} = 0$, then:

$$\mathbf{S} = |\mathbf{E}|^2 \mathbf{u}_k \tag{3.59}$$

The Poynting vector can be derived in our mathematical framework, from the wavevector and the electric field intensity. Let us define the angular scattering cross section at oblique incidence as: the energy flux radiated by the system per angle over a circular contour about the center of the defect. This can be written as:

$$\frac{\partial \sigma_{rad}}{\partial \alpha} = \rho\, S_\rho \tag{3.60}$$

where $S_\rho = \mathbf{S} \cdot \mathbf{u}_\rho$. Now, since:

$$S_\rho = |\mathbf{E}|^2 (\mathbf{u}_k \cdot \mathbf{u}_\rho) = |\mathbf{E}|^2 \ \frac{1}{g}(k_x \cos\alpha + k_z \sin\alpha) = \frac{\kappa}{g} \ |\mathbf{E}|^2, \tag{3.61}$$

given the far-field we previously calculated (in Eq.(3.45))

$$\mathbf{E}_{rad}(\rho,\alpha) = g^2 \frac{e^{i(\kappa\rho-\pi/4)}}{\sqrt{8\pi\kappa\rho}} \left[E^{(p)}_{rad}(\alpha)\mathbf{k}^+_p + E^{(s)}_{rad}(\alpha)\mathbf{k}_s \right] \tag{3.62}$$

we attain a clear-cut expression for the angular radiative emission as:

$$\frac{\partial\sigma_{rad}}{\partial\alpha} = \frac{g^3}{8\pi} \left[\, |E^{(p)}_{rad}(\alpha)|^2 + |E^{(s)}_{rad}(\alpha)|^2 \, \right] \tag{3.63}$$

The net energy radiated by the system is attained by summing up the energy radiated at each angle. Finally:

$$\sigma_{rad} = \frac{g^3}{8\pi} \int_0^\pi d\alpha \, \left[\, |E^{(p)}_{rad}(\alpha)|^2 + |E^{(s)}_{rad}(\alpha)|^2 \, \right]. \tag{3.64}$$

3.8 The Extinction Coefficient in 2D

The total scattering cross-section is defined as the sum of the scattering cross-sections of the inhomogeneity into both surface plasmon modes and radiative modes.

$$\sigma_{scat} = \sigma_{p+} + \sigma_{p-} + \sigma_{rad} \tag{3.65}$$

Let us only focus on the case of an incident surface plasmon. As proved in [37, 38] the total scattering cross-section coincides with the extinction cross-section in the case of a lossless medium, where the extinction cross-section is defined as:

$$\sigma_{xtn} = \frac{g}{S_{spp}} \int_A d\boldsymbol{\rho}' \, \mathbf{e}^*_{spp}(\boldsymbol{\rho}') \cdot \mathbf{P}(\boldsymbol{\rho}'). \tag{3.66}$$

Thus in the lossless case $\sigma_{xtn} = \sigma_{p+} + \sigma_{p-} + \sigma_{rad}$. However in general the extinction cross-section accounts for both the total scattering cross-section and the absorption of the inhomogeneity [37, 38], and thus $\sigma_{xtn} = \sigma_{scat} + \sigma_{abs} > \sigma_{scat}$, gives the total energy that is drawn out of the incident energy to be remitted and dissipated by the inhomogeneity.

3.9 Transmission Reflection and Radiation for 2D systems

Consider a bi-dimensional scattering systems illuminated by a surface plasmon polariton, $\mathbf{e}_{spp}(\mathbf{r})$ that propagates on a vacuum-metal interface, defined in Eq.(3.16) and Eq.(3.17). The impinging power from the SPP source on an inhomogeneity is broken up into three complementary streams of scattered Poynting vector flux. Some of the energy available goes into radiative loss out to free-space radiative modes (Radiation) which is $S = \sigma_{rad}$ of Eq.(3.64). The fraction of the scattered flux that relates to the SPP-mode propagation is well defined in terms of Transmission and Reflection of SPPs into other SPPs on the metal surface. The Reflection coincided with $R = \sigma_{p-}$ defined in Sec.3.4, as the emission of the system in a surface plasmon mode propagating in the direction opposite to the incident one, normalized to the plasmon energy flux S_{spp}. On the other side Transmission is defined as the fraction of the impinging energy flux that is neither scattered nor absorbed by the bi-dimensional inhomogeneity. Since all scattering cross-section are normalized to the impinging energy flux, the latter is equal to 1. Consequently the Transmission can be defined in terms of σ_{xtn} of Sec.3.8 and σ_{p+} of Sec.3.4 as:

$$T = 1 - \sigma_{xtn} + \sigma_{p+} \quad . \tag{3.67}$$

Namely we subtract from the impinging flux the net energy flux scattered and absorbed, and then return to it the energy flux scattered in the same direction as the incident flux.

Chapter 4

Device for launching surface plasmon polaritons

4.1 Introduction

Plasmonics is a branch of nanophotonics that is concerned with manipulation of surface plasmon-polaritons (SPPs) at nanoscale by means of specifically tailored metal structures. A great interest in this area is explained by two very interesting and promising applications for SPPs: integrated plasmonics and nanosensing. Integrated plasmonics, in contrast to integrated optics, operates with SPP waves, which can be squeezed much better than light, giving a possibility for high integration, and in turn transmit the signal much faster than electric current does. This provides a technology capable to replace electrical-circuit interconnections, strongly speed limited by the RC delay[39], with plasmonic ones limited merely by the SPP frequency[40, 41]. Moreover, plasmonics can bridge microscale photonics and nanoscale electronics[40]. Use of SPPs for nanosensing exploits surface plasmon resonances, which can be frequency shifted due to the presence of a substance being sensed as well as due to structure geometry. Furthermore, the latter, if chosen appropriately, can lead to an enormous field enhancement, increasing otherwise very weak Raman signal[42]. Both applications require high miniaturization of the components involved, and the issue of efficient *local* light coupling into SPP modes is an important concern. A very well known and widely used technique for excitation of SPPs is Kretschmann configuration. If a metal film illuminated from below is of appropriate thickness, and a parallel beam is used as a light source, almost all the incident power can be transmitted into plasmons, meaning that the efficiency of light to SPP coupling reaches nearly 100%. However, it is not a local excitation, since a parallel beam assumes its width to be at

least a few hundreds of microns, whereas plasmonic applications are in lack of efficient compact sources of SPPs of the several-wavelengths size. Moreover, the propagation length of an SPP beam along such a metal film is twice shorter than along an opaque one[43] due to the additional radiation losses into the substrate, which in most cases is undesirable. Otto configuration, which in common with Kretschmann one uses a similar principle of evanescent coupling, bears the same disadvantage of non-local excitation. A number of configurations have been developed for local light-SPP coupling[44]. Among them are excitation using illumination through a probe of a scanning near-field optical microscope[45], by a highly-focused laser beam[46, 47] or by a radially-polarized Bessel beam[48], excitation at discontinuities of a metal film[49], launching surface plasmons through an array of nano-holes[50], through a subwavelength slit supplemented with a periodic set of grooves[51], or by a surface protrusion defect in form of a particle or a single ridge[52], as well as by a periodic set of those[53]. The efficiency of light-SPP coupling, defined as the ratio of SPP power to that of light, for some configurations has been evaluated both experimentally and numerically[47, 48, 53–55]. The efficiency of the techniques that use radially-polarized Bessel or highly-focused parallel beams can be quite high[47, 48]. But apart from the already mentioned disadvantage of use of thin metal films, those methods, especially the former one, produce SPPs with virtually all directions of their k-vectors. This is very impractical for integrated plasmonics operating with SPP beams — laterally confined SPP waves with their k-vectors lying in a small range of directions (usually considerably below a radian) — but might be useful for nanosensing. In the work of Ditlbacher *et al.*[53], the efficiency of SPP excitation mediated by gold ridges has been studied. By illuminating the structures with a focused laser beam incident normally to the surface, two SPP beams propagating in opposite directions were excited. With only three ridges placed periodically on a gold film, the efficiency of light coupling into a single SPP beam was found to be $\sim 8\%$. This is already a very good result if one bears in mind that it is a local coupling configuration, and such commercial devices in serial production could be virtually fabricated on a microchip. A very recent study of light coupling to SPP for a single subwavelength hole in a gold film[54] revealed the efficiency of up to 28%. Note, though, that this is the value normalized with respect to the power incident onto the hole area, meaning that the absolute efficiency (normalized with respect to the total incident power) is considerably smaller. In our further discussions we shall operate with absolute values of efficiency. Note that a possibility of grating geometry optimization has been investigated numerically[55]. It was shown that a grating of indentations is more efficient than a grating of protrusions and an optimization of groove depth and width results in the coupling efficiency of 16%–22% to a single SPP beam. This chapter reports on the capability to optimize the ridge configuration for SPP

excitation, found already promising[53], with respect to its efficiency. We chose our structures to have a form of protrusions rather than indentations because ridges are easier to fabricate than grooves see[56]. We tried to optimize the geometrical parameters of an individual ridge (width and height), the number of those in a periodic array, as well as the working wavelength for a given array period. This chapter is not only based on our simulations of the systems, but also on their experimental verification with data provided by *S. I. Bozhevolnyi et al: Department of Physics and Nanotechnology, Aalborg University (Denmark).* What we achieved is a very satisfactory agreement between theory and experiment which constitute a tested device to launch surface plasmons in a single direction. The technical details of the experimental setup used by S. I. Bozhevolnyi et al, are not part of this thesis but can be found at [56] and in the literature cited below.

4.2 Local surface plasmon polariton excitation on ridges

We consider single ridges and periodic sets of those on top of a gold film illuminated with a focused laser beam. In the case of multiple ridges, the interaction between light and SPPs is achieved by coupling through the grating[43] and can be described approximately by the condition of momentum conservation:

$$\mathbf{k}_p = \mathbf{k} + \mathbf{u}_x \frac{2\pi}{\Lambda} n, \tag{4.1}$$

where $\mathbf{k}_p$ and $\mathbf{k}$ denote the wave-vector components in the (x, z) plane [see Fig. 4.1(a)] of respectively the scattered SPP and the incident light, Λ is the grating period, and n is integer. Since the grating momentum has only x-component, we consider light incidence in the (x, z) plane, hence $\mathbf{k} = \mathbf{u}_x g \sin\theta$ and Eq. (4.1) can be rewritten as:

$$k_p = g \sin\theta + nG, \tag{4.2}$$

with G denoting the grating momentum $2\pi/\Lambda$.

The number of ridges to be used for SPP excitation is changing from 1 (single ridge configuration) to at most 19. To find the exact optimum relation between the illumination wavelength and the grating period, it is enough to change only one of those parameters, and we choose to change the wavelength with the periodicity of ridges being fixed at $\Lambda = 800$ nm. We could change the height of the structures in our simulations but, because of experimental limitations[56], in the experiment all ridges have a fixed height. In this first Section it was set at 50-nm. To launch SPPs, we simulate a laser's Gaussian beam at normal incidence illuminating our structures with a tuneable (wavelength range 700–860 nm). The beam is focused to a spot with diameter estimated to be $\sim (5 \pm 0.5)$ μm (at the level $1/e^2$ of intensity). The air-gold interface is set in the (x, z) plane and the ridges are infinite along the y-axis, $k_y = 0$, and thus all the properties are invariant in this direction. If the incident electric field propagates in the (x, z) plane and has p-polarization [see Fig. 4.1(a)], then the total field and the scattered field will keep propagating in the (x, z) plane with the same polarization. That allows us to exclude the y-dependence in the calculations. For the incident electric field we assume (i) $E_y^{in} = 0$ and (ii) $E_x^{in}(x, z) = e^{-(x-x_0)^2/\sigma^2} E_x^{in}(z)$, i.e. the field has Gaussian profile only in the x-direction. Even though y-components of the incident electric field in the experiment are nonzero, with relatively weak focusing ($\sigma = 2.5$ μm) they are negligible[57]. Let the dielectric constants of gold (as well as the gold ridges) and air be respectively ε and 1. The system is formed by two semi-infinite half-spaces (air and gold) with embedded scatterers (ridges) at the interface. Since

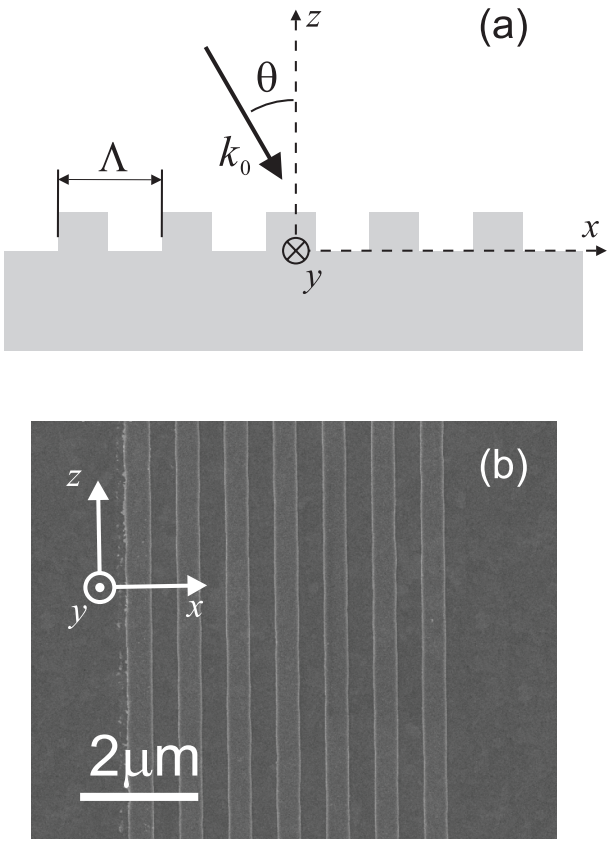

Figure 4.1: (a) Geometry of the configuration under study and coordinate system. (b) Scanning electron microscope image of a typical structure under investigation. In this case it is a set of 7 periodically arranged ridges. Period is 800 nm.

the structure under investigation is homogeneous along the y-direction, the polarization of the incident laser beam is perpendicular to the ridges. In this case two SPP beams are excited propagating in opposite directions, away from the structure. Note that in most application configurations only one of them can be used. For this reason we define the light-plasmon coupling efficiency as the ratio between the power carried by one of the SPP beams and that of the incident laser illumination, i.e. the percentage of power transferred from light into a SPP beam. In other words, we find a *unidirectional* SPP excitation efficiency. Note also, that with extended structures (like periodic set of ridges) the efficiency depends substantially on the position of the illumination spot. In this case we search for the position that gives the maximum efficiency. It is this unidirectional efficiency that we call the light-plasmon coupling efficiency or the SPP excitation efficiency. The random error of all measurements discussed in this work is estimated to be $\sim 15\%$.

4.3 Numerical Results and Experimental Agreement

4.3.1 Wavelength dependence of coupling efficiency

First, we investigated the wavelength dependence of the SPP excitation efficiency. For this purpose a 50-nm-high and 150-nm-wide single gold ridge was fabricated. A unidirectional SPP excitation efficiency was measured in the wavelength range 750–860 nm and found to decrease rapidly with the increasing wavelength. The most basic description of SPP launching in this configuration relies on a dipole scattering of light on the ridge. Its two sharp edges separated by 50 nm act as a dipole in the incident electric field oriented perpendicular to the ridge. Since the size of this dipole is considerably smaller than the wavelength, Rayleigh scattering gives an adequate description of the wavelength dependence of scattered light intensity[58]. We plotted the numerically obtained values along with the experimental data in a double-logarithmic chart to find the exponent in the power law dependence [Fig. 4.2(a)]. Both plots are fitted with a straight line, which yields the magnitudes of the assumed exponents -5.4 and -4.5 for the experimental and numerical data respectively. One can see that the agreement is fairly good both between each other and with Rayleigh scattering, which scales as a power of -4 of the wavelength. One should hardly expect the exact correspondence between the experimental and numerical data, because the simulated results depend much on the dielectric constant of metal (which is known to depend on the fabrication) as well as on the exact profile of the ridge, which is not ideally rectangular in the experiment.

In order to increase the efficiency, it seems reasonable to use several ridges arranged periodically. As mentioned in Section 4.2, the period should be close to the SPP wavelength, but the precise value is not obvious. We investigated the wavelength dependence of light-plasmon coupling efficiency for structures composed of up to 6 ridges having the same geometrical parameters as in the previous experiment [Fig. 4.2(b)]. Having the feature of a single-ridge dependence (to decrease rapidly with wavelength), the efficiency is noticeably enhanced near the wavelength of 790 nm where the coherently excited on different ridges SPP waves interfere with each other constructively, increasing the light-plasmon coupling. Note that while the number of ridges used is increasing, this effect starts to be more pronounced. For one ridge we obtained a unidirectional efficiency around 0.6% at the wavelength 800 nm, which is almost four times less than that declared in Ref. [53]. This is due to the difference in laser illumination spot size (~ 5 μm in our case vs. 1 μm in Ref. [53]). Although better focusing is preferred when the system consists of a few ridges, a larger beam size might be more advantageous for a considerable number of

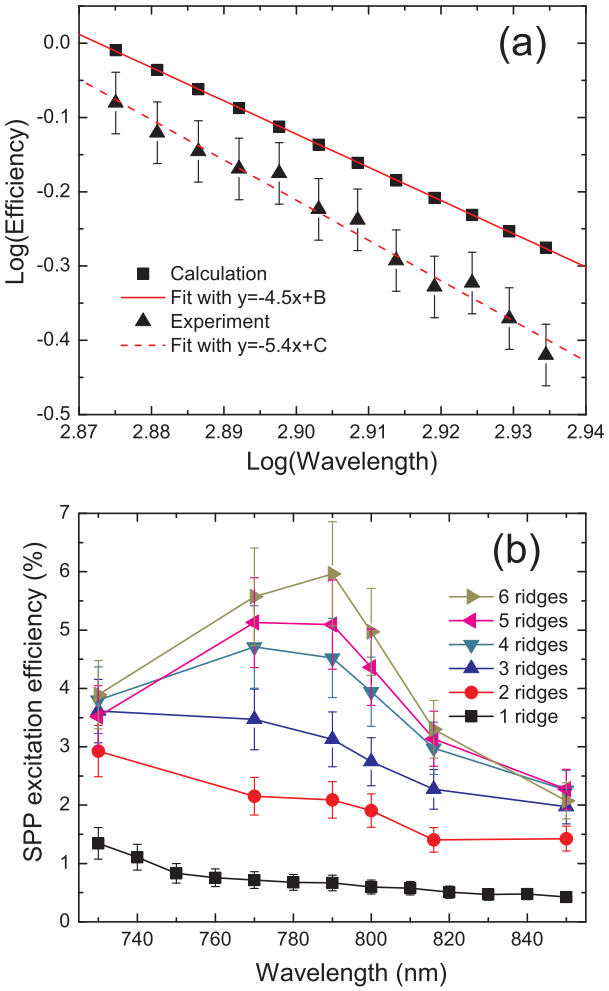

Figure 4.2: (Colour online) (a) Efficiency of SPP excitation on 50-nm-high and 150-nm-wide gold ridge versus the free-space wavelength used for excitation plotted in double-logarithmic scale along with a linear fit. (b) Efficiency of SPP excitation on sets of periodically arranged 50-nm-high 150-nm-wide gold ridges versus the free-space wavelength measured for the different number of ridges composing the grating. Separation between ridges is $\Lambda = 800$ nm.

ridges (10–20) since more scatterers are involved in the excitation process.

4.3.2 Dependence on geometrical parameters of ridge

Finding the optimum geometrical parameters of an individual ridge is important because, as we show in this Section, geometrical optimization can sufficiently increase the efficiency of light-plasmon coupling. We performed numerical simulations of SPP excitation on a single ridge whose height is varying in the range 50–230 nm and width — in the range 200–600 nm, and calculated its efficiency [Fig. 4.3(a)]. With only small difference, ridges of all heights feature maximum efficiency when the width is close to 350 nm, i.e. almost half of the wavelength. The maximum achievable efficiency for every given height is then plotted in Fig. 4.3(b). One can see almost a 6-fold increase of the efficiency with the height growing from 50 nm up to 130 nm, where it reaches its maximum.

The numerical results for the width dependence of excitation efficiency were checked experimentally (by I.P.Radko et al.), also using illumination at the wavelength 800 nm. The ridge height was fixed at 50 nm. They tested configurations with 1, 3 and 5 ridges aligned periodically (period $\Lambda = 800$ nm) in the two latter cases [Fig. 4.3(c)]. The experimental data exhibit maximum efficiency for the ridge width close to 350 nm, which is in agreement with the numerical simulations. The figure also demonstrates that the optimum width is not changing for the increasing number of ridges, at least within a small quantity. Note that the fact that we obtained larger efficiency experimentally, than it is predicted numerically for a 50-nm-high ridge, can be explained by a slightly smaller laser beam diameter used in the experiment.

4.3.3 Optimum wavelength for excitation on ridges

As we already mentioned, one of the way to increase the efficiency of excitation is to use a periodic set of ridges. However, the optimum period is not necessarily equal to the SPP (or free-space) wavelength. To get more insight into that, we investigate arrays of ridges with the fixed period of $\Lambda = 800$ nm and change the laser beam wavelength instead, since array period cannot be changed in the experiment. First, we investigated numerically SPP excitation on a grating composed of 11 ridges having width 280 nm and height 50 nm. The diameter of the modelled laser beam is 5 μm. We placed the zero of the coordinate system to the centre of the grating [inset in Fig. 4.4(a)]. Then we performed a scan of the laser beam from the centre towards increasing coordinates, evaluating the powers carried by the two SPP beams propagating in opposite directions for different wavelengths [Fig. 4.4(a,b)]. Expectedly, the efficiency drops down

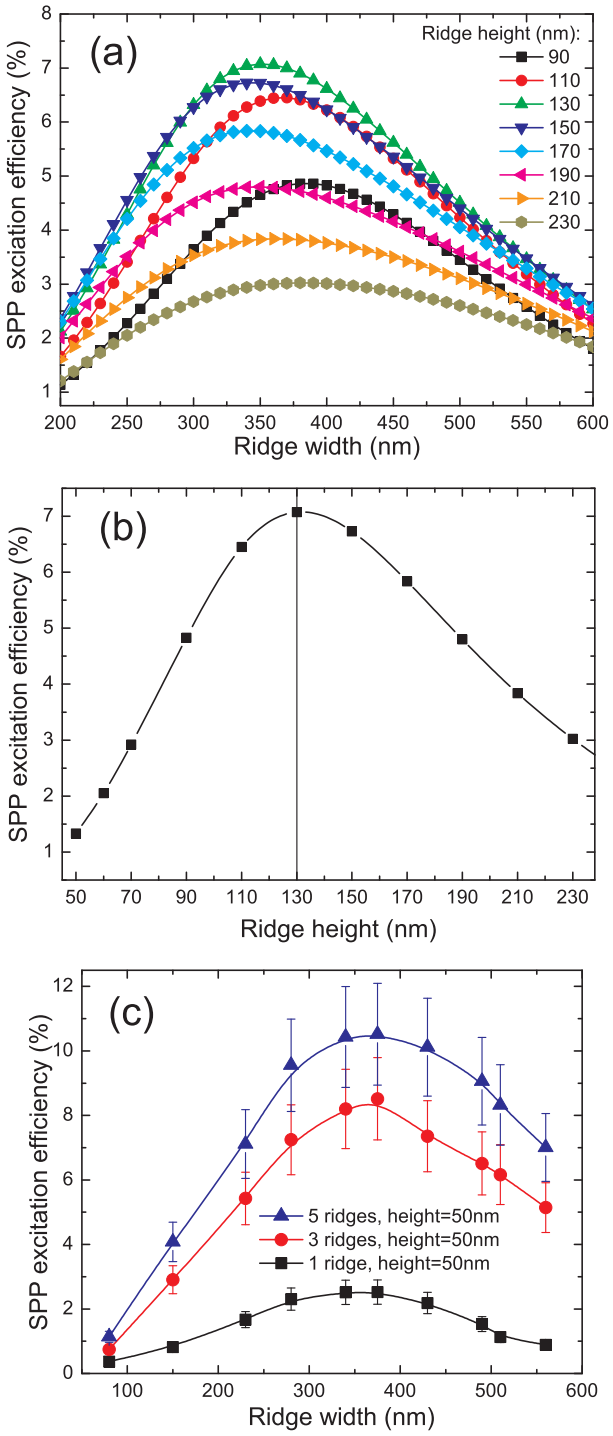

Figure 4.3: (Colour online) (a) Calculation of efficiency of SPP excitation on a single ridge versus its width. Different curves shows data for different ridge heights. (b) Calculation of maximum attainable efficiency (with all possible ridge widths) of the SPP excitation on a single ridge versus its height. (c) Experimental results for dependencies of SPP excitation efficiency on the ridge width. The ridge height is fixed at 50 nm. Three curves show three sets of measurements accomplished on a single ridge and on 3 and 5 ridges aligned in gratings with the period $\Lambda = 800$ nm. The free-space wavelength is 800 nm both in calculations and in the experiment.

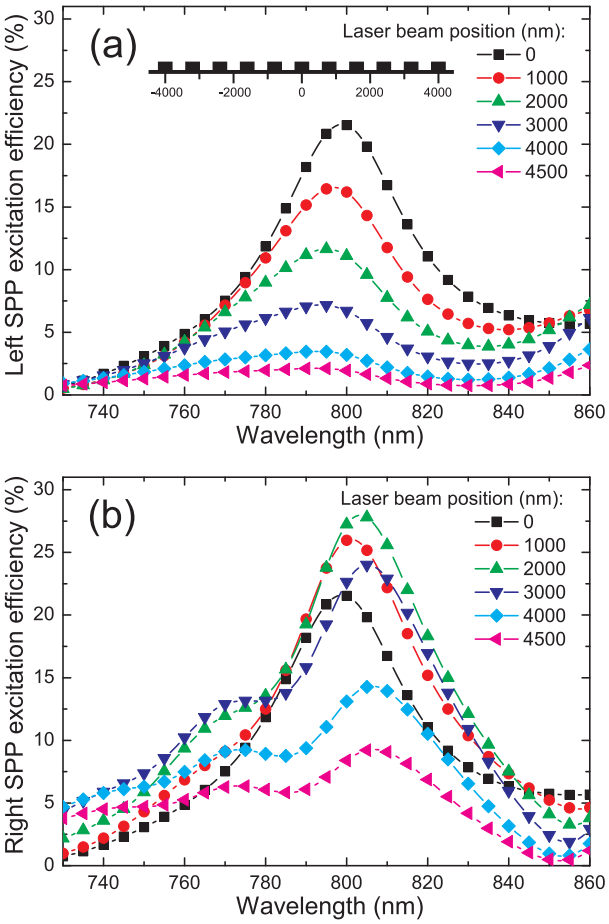

Figure 4.4: (Colour online) Numerical results for wavelength dependence of efficiency of SPP excitation on 11 ridges (height = 50 nm, width = 280 nm). Different positions of a laser beam (diameter = 5 μm) illuminating the sample are shown by different curves. Calculations accomplished for SPP propagating to the left (a) and to the right (b) from the grating. Inset in the panel (a) shows the geometry and coordinate system.

monotonously for the left-propagating SPP beam [see the inset in Fig. 4.4(a) for the geometry] with the laser beam moving from the centre of the grating to the right [Fig. 4.4(a)]. On the contrary, the efficiency for the right-propagating SPP increases first, reaching the maximum at the coordinate ~ 2000 nm, and then monotonously decreases [Fig. 4.4(b)]. We will get back to this behaviour in the next Section, whereas now a more important result is the wavelength where the maximum efficiency is observed: it is 805 nm. Note that in this optimum case, the grating period of 800 nm is between the free-space wavelength, 805 nm, and the corresponding SPP wavelength, ~ 789 nm. Then we looked into the experimental data, where they investigated two gratings composed of 7 and 11 ridges of the width 200 nm and 280 nm each, respectively. Instead of scanning the laser beam across the gratings, they were searching the position with the maximum achievable efficiency and then picked that efficiency value to the chart. Thus, the way it was measured gives an envelope of all the curves in Fig. 4.4(b). The result obtained is shown in Fig. 4.5(a) together with the envelope for comparison. The lower curve yields the optimum wavelength of 790 nm, which repeats the result shown in Fig. 4.2(b). The experimental curve for the 11-ridges grating shows that the optimum wavelength shifts towards higher values with the increasing number of ridges. Comparison with the envelope of the curves in Fig. 4.4(b) (obtained for the same geometry, i.e. 11 ridges of the width 280 nm and height 50 nm) shows good agreement in both the optimum wavelength and the general behavior of the spectrum. For lower wavelengths the experiment gives higher efficiencies. This is because the experimental structure always has small irregularities in the ridge periodicity, which makes the spectrum broader. A drift of the optimum wavelength with the increasing number of ridges can be also observed in another series of measurements where we investigated a dependence of light-SPP coupling efficiency on the number of ridges involved in the process [Fig. 4.5(b)]. The ridges participating in the excitation are 280-nm-wide. First of all, at wavelengths far from the optimum one, the efficiency comes very quickly to saturation (730 nm and 770 nm) or even drops down (850 nm) with the number of ridges exceeding 5. Secondly, for the small number of ridges (below 7) the optimum wavelength is somewhere around 800 nm, but cannot be clearly defined: the gratings are broadband. However, with the increasing number of ridges, the wavelength dependence gets more pronounced: first, 800 nm is preferable (for 11 ridges) and then 810 nm (for 15 and more ridges). The two main conclusions are: (i) the optimum wavelength (or the optimum grating period for a given wavelength) is growing (diminishing) with the increasing number of ridges composing the grating, and (ii) for a grating with large number of ridges the SPP wavelength corresponding to the optimum illumination is approaching the grating periodicity from the low-wavelength side. The latter means that there are two competing processes of electromagnetic interaction between ridges: one is on

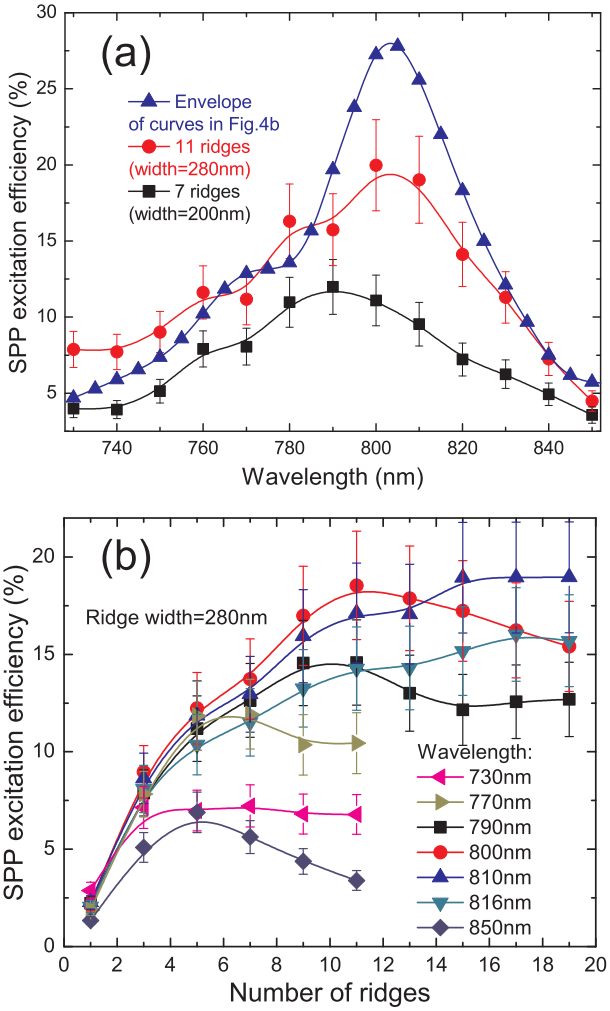

Figure 4.5: (Colour online) (a) Experimental results for wavelength dependence of efficiency of SPP excitation on two different gratings ($\Lambda = 800$ nm) composed of 7 ridges (height = 50 nm, width = 200 nm) and 11 ridges (height = 50 nm, width = 280 nm) plotted together with the envelope of curves shown in Fig. 4.4(b) for comparison. (b) Experimental results for efficiency of SPP excitation versus the number of ridges in the grating ($\Lambda = 800$ nm) used for excitation. Ridge height = 50 nm, width = 280 nm. Results obtained at different wavelengths are shown by different curves.

the dielectric (air) side and mediated by photons, and the other is on the metal side occurring through SPPs. To enhance each of those processes, the grating period equal to the corresponding wavelength is required, which sets the optimum period in between of them. As it turns out, however, the interaction through plasmons is stronger, and with the large number of ridges it takes over the process, equating the optimum period with the SPP wavelength. One of the reasons that we obtained the maximum efficiency at $\lambda_0 = 810$ nm instead of 816 nm (corresponding to $\lambda_{\mathrm{spp}} = \Lambda = 800$ nm) might be a small misalignment (of about 0.5°) of the objective focusing illumination onto the grating in the experimental work [see Eq. (4.2)].

4.3.4 Directionality of SPP excitation on periodic sets of ridges

We continue our consideration of efficiency of light-SPP coupling on ridges by demonstrating the extent of asymmetry of surface plasmon beams excitation while using such structures. This gives an idea of the total (bidirectional) light-plasmon coupling efficiency and also demonstrates the sensitivity of efficiency to the position of the laser beam with respect to the illuminated structure. For this purpose, we calculate the efficiency while the laser beam is scanned across a grating [Fig. 4.6(a)]. Such profiles were recorded at several wavelengths. The grating was composed of 11 ridges (height = 50 nm, width = 280 nm). At two wavelengths — at which the highest efficiencies were obtained (800 nm and 816 nm) — the profiles feature a relatively sharp maximum. At the optimum wavelength the efficiency is very sensitive to the laser beam position, which leads to a more unidirectional excitation. Assuming that the centre of the grating is approximately positioned at $x = 7$ μm, one can find the efficiency of SPP excitation on the farthest end of the grating while having maximum at the nearest one: 7% versus 19% at $\lambda_0 = 800$ nm. For the wavelengths out of the resonance, the profiles possess a small plateau in the middle forming a shape reminding trapezium. At the wavelength 730 nm one can even notice a very small second local maximum on the other end of the grating.
Modeling of this geometry shows an excellent agreement [Fig. 4.6(b)] the shape of the curves, including the two local maxima at $\lambda_0 = 730$ nm. The explanation for the latter could be that, out of the resonance, the excitation on fewer number of ridges is more efficient than on the whole grating because the periodic set of ridges causes destructive interference. Therefore, moving the beam out of the grating excludes most of the ridges from the process, which turns out to be more favorable in this case.

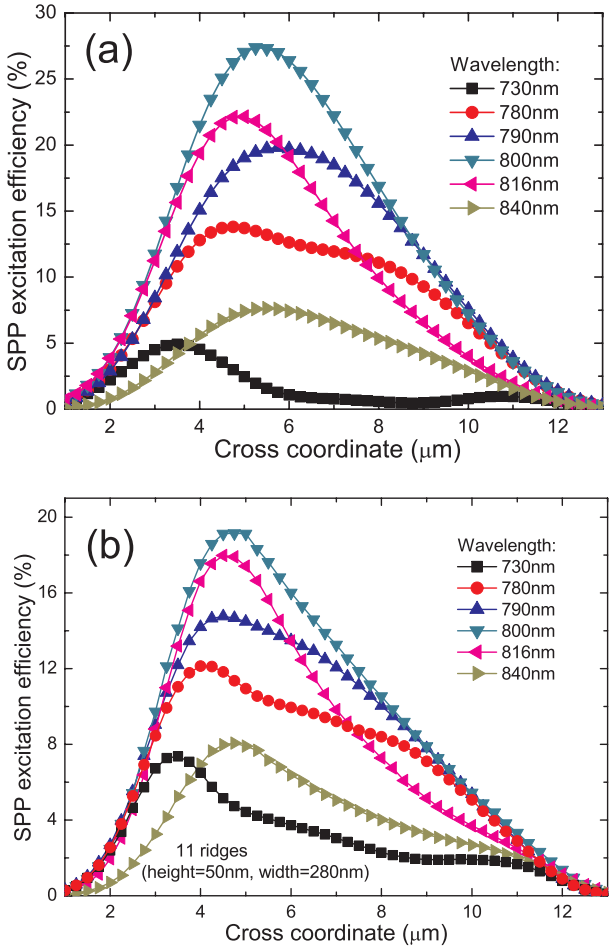

Figure 4.6: (Color online) Numerical (a) and experimental (b) results for SPP excitation efficiency versus the position of a laser beam scanned across the grating (Λ = 800 nm) composed of 11 ridges (height = 50 nm, width = 280 nm). In the experiment the grating extends approximately from 3 to 11 μm of the cross coordinate, in the numerical simulations this is the exact position of the grating. The error bars in (b) are omitted to not obscure the image.

4.4 Efficient unidirectional ridge excitation of surface plasmons

We have showed the possibility to efficiently convert a focused laser beam (at normal incidence) into an SPP beam (i.e., a laterally confined SPP wave possessing small divergence) using periodic set of metal ridges on top of a metal surface [51]. Using 50-nm-high and 280-nm-wide gold ridges we have achieved

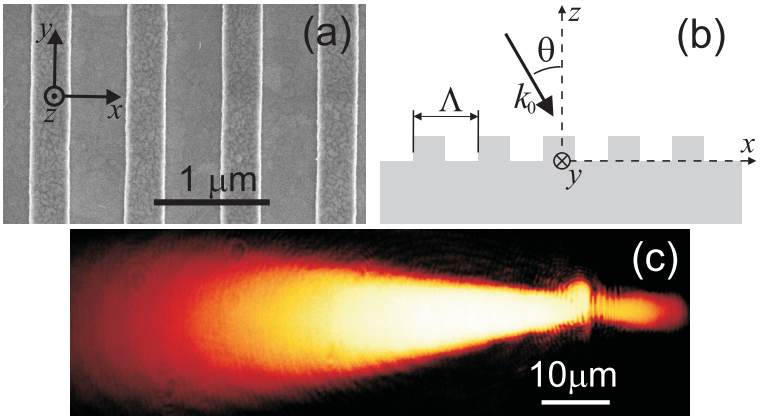

Figure 4.7: (a) SEM image of a fragment of the fabricated structure consisting of 130-nm-high and 330-nm-wide ridges. (b) Geometry of the illumination configuration. (c) Typical LRM image of a strong SPP beam excited on the 11-ridge array on its left side. A much weaker SPP beam propagating to the right is also visible.

the SPP excitation efficiency (defined as the power ratio between an SPP beam propagating in a given direction and an incident laser beam) of ~ 0.2 at the wavelength of ~ 800 nm, which we believe is the best result obtained so far. We have also shown via numerical simulations that one can dramatically increase this efficiency by propitiously choosing the ridge dimensions found for this configuration to be ~ 130 nm in height and ~ 350 nm in width. Now we shall investigate how these ridge SPP couplers can be optimized by changing parameters. The configuration for SPP excitation exploited in this work is the same but we change the defect sizes. This time we consider straight 130-nm-high and 330-nm-wide gold ridges (of nominally rectangular profile). The number of ridges in the configuration under investigation changes from one (single-ridge configuration) to 15 in steps of two with the period being fixed at $\Lambda = 800$ nm [Fig. 4.7(a)]. The illumination is kept the same, a spot with a diameter estimated to be $\sim (5 \pm 0.5)$ μm (at the level $1/e^2$ of intensity) linearly polarized in the direction perpendicular to the ridges [which run parallel to the y axis in Fig. 4.7(b)] resulting in the excitation of two SPP beams propagating in opposite directions, whose intensities strongly depended on the position of the illumination spot relative to the structure [Fig. 4.7(c)]. Note that unless

otherwise stated, the SPP excitation efficiency refers to the SPP beam propagating to the left [i.e., towards negative x coordinates in Fig. 4.7(b)] from the structure.

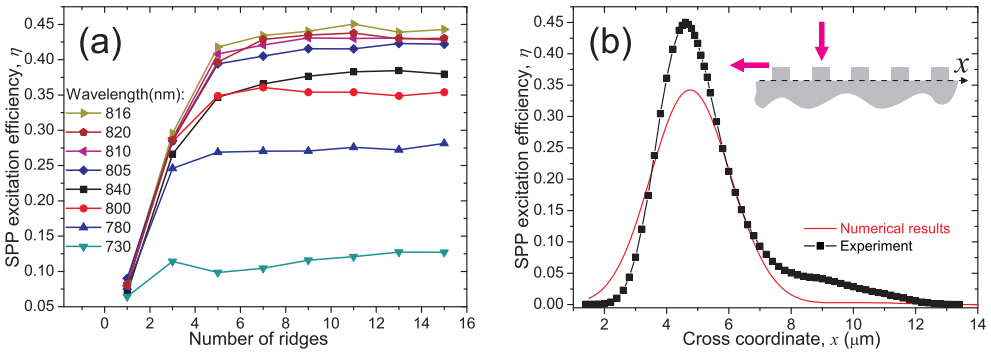

Figure 4.8: (a) The maximum efficiency of SPP excitation obtained experimentally for different laser wavelengths as a function of number of ridges. The lines connecting the points are only to guide an eye. (b) Experimental and numerical results for the SPP excitation efficiency versus the position of a laser beam scanned across the 11-ridge array extending from $x = 4$ μm to $x = 12$ μm. Error bars both in (a) and (b) are omitted for clarity of presentation. The error of measurements is estimated to be $\sim 13\%$.

Efficient SPP excitation (in the first grating order) at normal incidence requires thereby that the grating period should be equal to the SPP wavelength: $\Lambda = \lambda_{\mathrm{SPP}}$. In this case, the laser illumination scattered on different ridges generates coherent SPP waves, which interfere constructively with each other increasing the efficiency of SPP excitation.

We started our work by analyzing the experimental data received by S. I. Bozhevolnyi et al. They used a discrete set of the laser (free-space) wavelengths centered at the wavelength of 816 nm corresponding to $\lambda_{\mathrm{SPP}} = \Lambda = 800$ nm [43]. For each wavelength and number of ridges, the incident laser beam was laterally adjusted [along x axis in Fig. 4.7(b)] so as to maximize the excited SPP beam power [Fig. 4.8(a)]. It is seen that the dependence of the SPP excitation efficiency on the number of ridges exhibits a rapid saturation [Fig. 4.8(a)], even more rapid than in the case of the 50-nm-high ridges (cf. with Fig. 4.5(b) of the previous Section). The highest efficiencies were measured at the free-space wavelength of 816 nm as expected. At this wavelength, the SPP excitation efficiency of 0.45 ± 0.06 was achieved with 11 ridges, reaching the level of > 0.4 already with five of them. The excitation efficiency is very sensitive to the transverse position of laser beam relative to the grating [51], rendering the (practically) unidirectional SPP excitation [Fig. 4.7(c)].

We simulated the efficiency measurements numerically. The laser beam was scanned across the 11-ridge array illuminated at the wavelength of 816 nm

[Fig. 4.8(b)], both in the simulation and in the experiments. Taking into account that the array center is positioned at $x = 8$ μm, one can deduce the excitation efficiency, η, for the (unwanted) SPP beam propagating to the right when the left-propagating SPP beam is most efficiently excited. Since the highest efficiency of 0.45 is achieved at $x = 4.6$ μm, the corresponding efficiency for the right-propagating SPP beam should be (under the assumption of symmetric excitation configuration) the same as that for the left-propagating one at $x = 8 + (8 - 4.6) = 11.4$ μm, i.e. $\eta = 0.013$, which is ~ 35 times smaller than 0.45 [Fig. 4.8(b)]. We think that such an efficient suppression of the right-propagating SPP beam can be due to the second-order Bragg reflection from ridges, which redirects the SPP back into the desirable direction, as well as due to the (reciprocal) out-coupling of the propagating (over many ridges) SPP, which can be rather strong since the in-coupling is efficient.

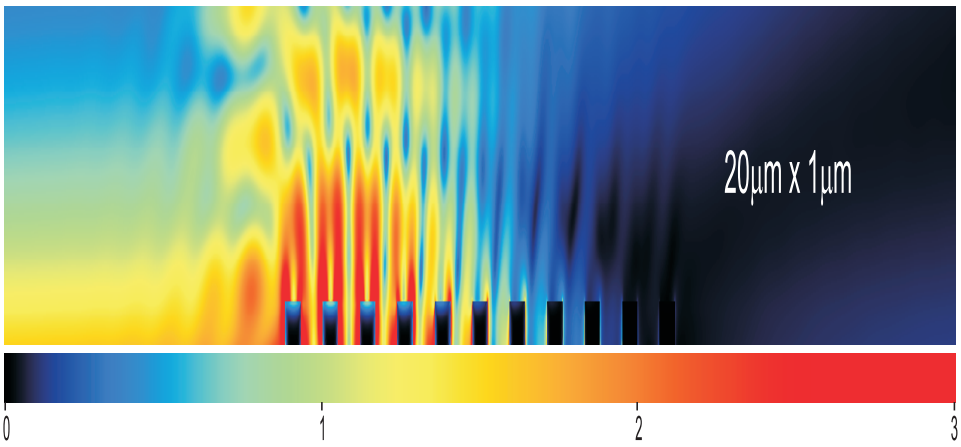

Figure 4.9: Numerical results for electric-field amplitude distribution near the 11-ridge array illuminated under the optimal conditions for SPP excitation. The field values are normalized to the amplitude of the incident field.

We calculated in Sec.4.2 the corresponding dependence of SPP excitation for the nominal structure parameters [Fig. 4.8(b)]. There is a good agreement between the experimental and simulated dependencies in their shape, albeit the simulated SPP excitation is more unidirectional than that observed in the experiment, ensuring ~ 140 times more efficient SPP excitation to the left than that to the right. The latter is believed to be due to perfect geometry used in the numerical simulations, featuring identical ridges with exactly the same separation, whereas fluctuations in the geometrical parameters are expected in the experiment. There is also a discrepancy in the maximum value of efficiency, which does not fall into the interval determined by measurement errors. This difference can be explained by deviations in the ridge parameters (in ridge shape and sizes) of the fabricated structure from the nominal values. It is also

known [59] that the dielectric constant of gold within ridges may be different from that reported in literature [60, 61].
Having realized such a high efficiency of (local) SPP excitation, we anticipated to also achieve the field enhancement near the gold surface with respect to the field of the incident laser beam. Continuing our simulations of the 11-ridge array, we calculated the electric-field amplitude (normalized to the amplitude of the incident laser beam) distribution near the surface for the maximum SPP excitation (Fig. 4.9). Unidirectional excitation of the left propagating SPP is clearly displayed in this distribution showing also that the electric field is indeed enhanced by ∼1.5–1.9 times on the left side of the array due to the SPP excitation. One should keep in mind that, upon direct illumination of a smooth metal surface, the field at the surface is rather small due to the reflectivity being close to -1. In this respect, the obtained field enhancement is quite an achievement and very promising for sensing applications, since the enhanced (propagating) field covers a large surface area [Fig. 4.7(c)] thus allowing the interrogation of a considerable amount of a test substance.

4.5 Conclusions

Summarizing, first of all in Sec.4.3 we have investigated theoretically the efficiency of SPP excitation on single ridges and on periodic sets of those and supported our results experimentally. The variable parameters were wavelength of light used for excitation, ridges width and height, as well as the number of them in periodic arrangements. The maximum theoretical unidirectional SPP excitation efficiency of about 28% was obtained for 11 ridges (width = 280 nm, height = 50 nm). Our numerical simulations show that the dependence of efficiency on ridge height is an important parameter, which can lead to almost 6-fold increase in efficiency for one ridge when changing the height from 50 to 130 nm. The optimum ridge width was found to be close to half the illumination wavelength, but somewhat smaller. This is in agreement with the result obtained previously for another working wavelength[55]. For periodic sets of ridges used for SPP excitation, the SPP wavelength corresponding to the optimum free-space wavelength was found to be smaller than the ridge periodicity, but approaching that with the increasing number of ridges.
It is difficult to extend the conclusions to systems distinct from the suggested one without thorough investigations. One can, however, conclude that efficient local excitation of SPPs is achievable with a moderately focused laser beam and only a few ridges. Regarding optimum parameters, it is clear that with the oblique incidence of the laser beam, the grating period and/or the excitation wavelength have to be adjusted according to Eq. (4.2). It is further expected that the ridge parameters (height and width) are mainly influenced

by the choice of wavelength (most probably scaling with it linearly), but the precise values have to be found with detailed simulations. Finally, we note that the value of efficiency of 20% found experimentally might be improved in the case of opaque metal film[55], which prevents the excitation of SPPs on a metal-substrate interface. The higher value of $\sim 28\%$ obtained numerically for the same configuration is a good evidence of that. However, the use of opaque metal films will make the LR microscopy inapplicable for experimental investigations.

Finally in Sec.4.4 we have characterized the efficiency of unidirectional surface-plasmon excitation with periodic (800 nm) arrays of 130-nm-high and 330-nm-wide gold ridges on a thin gold film illuminated with a focused (5-μm-wide) laser beam. We have demonstrated that, at the resonant wavelength of 816 nm, the excitation efficiency of > 0.4 can be obtained with ≥ 5 ridges by adjusting the beam position. Conducting numerical simulations we have calculated the electric-field enhancement achieved near the gold surface. We have demonstrated an efficient unidirectional SPP excitation with a focused laser beam achieving the SPP excitation efficiency of 0.45 ± 0.06 at the resonant wavelength of 816 nm. Numerical simulations and the experimental results are in agreement and provide an estimate of the field enhancement obtained near the gold surface due to the SPP excitation. The presented approach to generation of enhanced electric fields may be an alternative to localized plasmons for achieving enhanced fluorescence and Raman scattering [62].

Chapter 5

Scattering of surface plasmon polaritons by 2D impedance barrier

5.1 Introduction

Recent advances of nanotechnologies have disclosed a wide prospect for designing and realizing novel optical devices with applications in information processing and communication. Surface plasmon polariton (SPP) resonances could be employed to improve the efficiency of photonic circuits by increasing, for instance, the output power of surface emitting diodes, [63] or decreasing the size of optical waveguides to a subwavelength scale.[64]
When metallic elements are used, the high localization of electromagnetic field associated to SPPs can be exploited to guide light into volumes significantly smaller than the diffraction limit. But, the successful control of SPPs requires the implementation of optical elements for these surface modes such as mirrors, light-emitters, multiplexers and so forth. [65–67]
It has been demonstrated, both theoretically and experimentally, that even shallow subwavelength surface defects or nanoparticles significantly scatter the incident energy from a SPP.[68, 69]
Interestingly, the difference between the out-of-plane scattering cross-sections for relief and impedance inhomogeneities is considerable, even when these inhomogeneities have the same spatial dependency.[70]
The aim of this chapter is to examine the scattering of SPPs by one-dimensional (1D) defects of the surface impedance when the SPP angle of incidence is varied. Since it is much easier to measure the reflection coefficient at oblique incidence[69] than it is at normal incidence,[71] an investigation of the angle-

dependent scattering amplitudes is relevant to possible experiments.

5.2 Scattering of surface plasmon polaritons by impedance barriers: dependence on angle of incidence

5.2.1 The theoretical methods

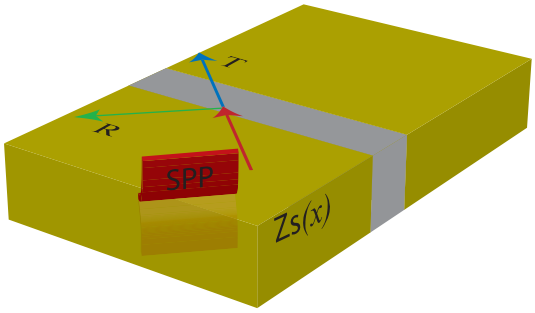

Figure 5.1: Schematic illustration of the studied system: SPP scattering at the inhomogeneity formed by the perturbation of the surface impedance $Z_s(x)$.

Consider a SPP, $\mathbf{e}_{spp}(\mathbf{r})$, defined in Eq.(3.16) and eq.(3.17) that propagates on a vacuum-metal interface and impinges at angle θ onto a one-dimensional (1D) perturbation of the surface impedance, $Z_s(x)$. In terms of the dielectric permittivity of the metal, ε, the surface impedance is defined as $Z_s = 1/\sqrt{\varepsilon}$, see Fig. 5.1.
In this section the results of the Green Tensor Approach will be contrasted against those obtained by *A. Yu. Nikitin* with a different theoretical method. The Green Tensor Approach gives a virtually exact result and the method developed by A.Yu. Nikitin is an approximation, which provides analytical results. While the details of such method are not part of this thesis and are reported elsewhere[36], we shall mention exclusively the aspects of the approximations involved in his theoretical framework, that are functional to the discussing the agreement between the two methods. One approximation consists in substituting the electric and magnetic field boundary conditions at the interface with Surface Impedance Boundary Conditions(SIBCs). This approximation only works when $|\varepsilon| \gg 1$, and $|Z_s| \ll 1$. The other relevant approximation is the First Born Approximation (FOBA). This consists in approximating the field in the source region of the defect A, with the incident field $\mathbf{e}_{spp}(\boldsymbol{\rho}')$.
Absorption is expected to be negligible within the region of the defect, provided the width of the impedance defect is much smaller than the SPP propagation length. Therefore the metal is assumed to be lossless along the SPP path. We

will enforce this simplification by using only the real part of tabulated dielectric constants of the metals and setting $\mathrm{Im}(\varepsilon) = 0$. Nevertheless, in what follows, we will also show how absorption would affect the results.

5.2.2 The scattering system

Let us focus on inhomogeneities having a rectangular shape, with the following x-dependency of the impedance

$$Z_s(x) = \begin{cases} Z_s, & |x| > a/2\,; \\ Z_{si}, & |x| \le a/2\,. \end{cases} \tag{5.1}$$

Such defects may be manufactured by inserting a metal wire, of rectangular cross-section and impedance Z_{si}, into a thick conducting film with impedance Z_s. Although the considered impedance defect is infinite along both the y and z axis, it can be represented by a two-dimensional inhomogeneity of the metal dielectric permittivity with a finite depth in coordinate z, as long as this depth is much greater than the skin-depth of both the background metal and the defect. For example, we find that the scattering coefficients of an aluminium defect with depth 30 nm in the silver half-space differs only by 2% with respect to the infinite depth case (for the defect width 200 nm and $\lambda = 600$ nm). Such depth is ≈ 2 skin-depths in aluminum. In all GTA calculations shown, a defect of depth $z_\infty = 40$ nm was used to simulate a defect of infinite depth.

5.2.3 Solutions and Results: The surface plasmons Brewster angle analog

We have contrasted the exact but numerical Green's tensor calculations with the A.Yu.Nikitin solutions, which have the great advantage of being analytical. Even so we can get the same solutions as Nikitin's, by applying in framework of the Green tensor **just one** of his approximations: the First Born Approximation. If the polarization induced in the defect is taken to as:

$$\mathbf{P}(\boldsymbol{\rho}') = \Delta\varepsilon \; \mathbf{e}_{spp}(\boldsymbol{\rho}') \tag{5.2}$$

then the SPP reflection coefficient can be obtained by eq.(3.30) as:

$$R = |E_{p-}|^2 \tag{5.3}$$

where by plugging eq.(5.2) into eq.(3.32) we end up with:

$$E_{p-} = \frac{\Delta\varepsilon}{\varepsilon} \left(\sin^2\theta - \cos^2\theta\right) \; \frac{e^{\mathtt{i}k_{px}a} - 1}{\mathtt{i}k_{px}} \; \frac{e^{-|k_{pz}|\, z_\infty} - 1}{\mathtt{i}k_{pz}} \tag{5.4}$$

Eq.(5.4) yields two types of zeroes in the reflection coefficients, which appear as deep minima in the reflectance when going beyond the FOBA. The first zero of reflection arises from the vanishing of R. It appears at angle of incidence $\theta_B = \pi/4$ and is independent of the size of the defect. This reflection minimum for SPPs is reminiscent of the zero reflectance at the Brewster angle appearing when a p-polarized wave impinges onto a dielectric interface. Its existence can be understood, mutatis mutandis, following the explanation for the appearance of the Brewster angle. Consider a SPP impinging on an impedance defect. Since the impedance defect is located in the metal half-space, it can be represented by a polarization directed along the electric field of the SPP *inside the metal*. The field of the SPP inside the metal has a predominant longitudinal component parallel to the surface, with the z component being smaller by a factor $1/\sqrt{\varepsilon}$ (see eq.(3.33) of Sec.3.4.3). Therefore the polarization of an impedance barrier will point mainly in the direction of incidence. As shown in Sec.3.4.3, a dipole parallel to the surface emits SPPs primordially in the longitudinal direction, and not at all in the perpendicular direction. At $\theta_B = \pi/4$ the perpendicular direction coincides with the direction of reflection, so the reflection coefficient vanishes. For relief defects, however, the polarization mainly points along the direction normal to the surface as the incident SPP has a predominant z-component above the metal. The radiation of SPPs by a point dipole in the vacuum half-space directed in the z-direction is isotropically symmetric. This explains why, in this case, the reflection coefficient for an interface, R, does not depend on the angle of incidence (see Sec.3.4.3).

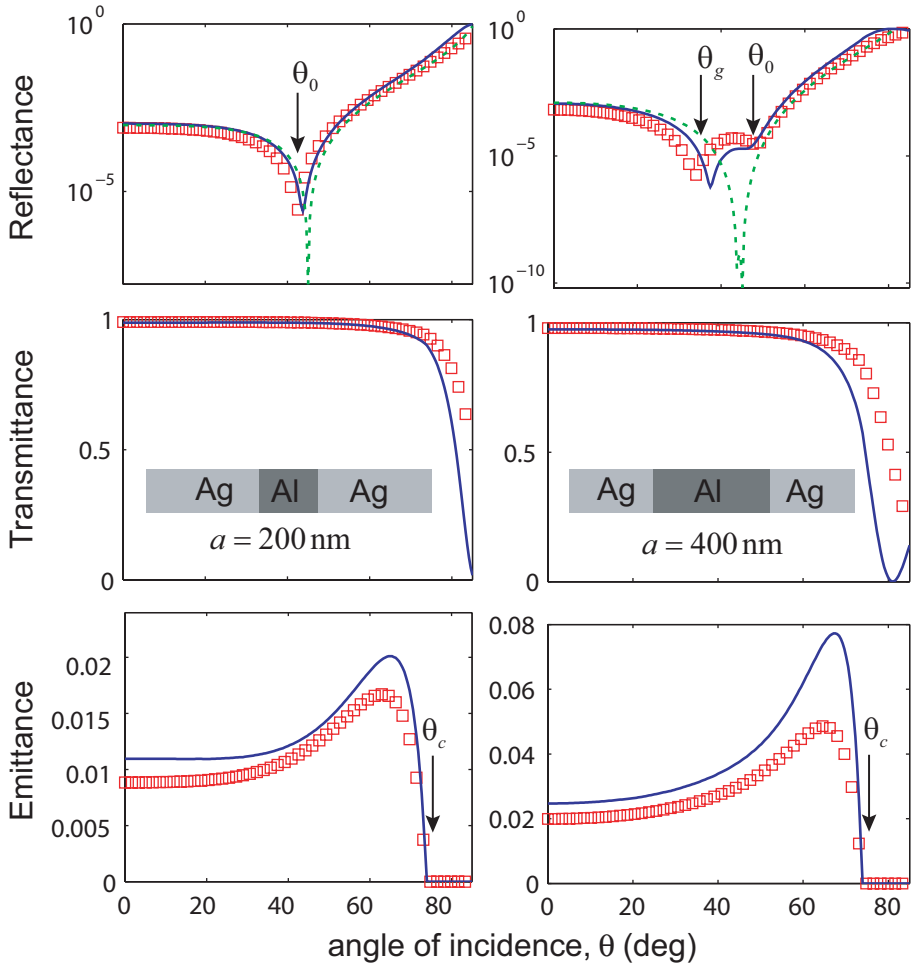

Figure 5.2: (Color online) The SPP reflectance R, transmittance T, and emittance S as functions of the incident angle θ. SPP impinges onto an aluminum stripe of rectangular cross-section placed in a thick silver slab. The defect widths are of 200 nm on the left-hand panel and 400 nm on the right-hand panel. Results obtained within the Rayleigh approximation are rendered by solid curves (full calculation) and dashed lines (FOBA). Squares represent the results of the GTA.

The second type of reflection minima is given by the condition $k_{px}a = n\pi$, namely $2q_p(a/\lambda)\cos\theta_g = n$, where n is an integer and $q_p = k_p/g$. This reflectance minima is due to interference: it occurs when the optical path of the plasmon inside the barrier is such that the amplitudes of reflected SPPs arising from the two ends of the barrier are in anti-phase. Notice that such optical path depends on both the angle of incidence and barrier width. Fig. 5.2 represents the scattering coefficients for a SPP on an air-silver interface, impinging onto an aluminum defect, as a function of angle of incidence. The left-hand and right-hand panels show respectively the results for a defect having a width of

200 nm and 400 nm. All dependencies have been calculated at $\lambda = 600$ nm taking ε from Ref. [61] [$Z_{s\,Ag} = -0.277i$, $Z_{s\,Al} = -0.146i$]. The squares represent the results of the GTA while the solid curves were computed by A.Nikitin. In this calculation the impedance obtained within the framework of SIBCs, was adjusted to $Z_s \to 1/\sqrt{1+\varepsilon}$. With this minor change the expression for q_p within the SIBCs coincides with the exact one. Both cases considered in Fig. 5.2 exhibit the Brewster-like reflectance minima at $\theta = \theta_B = 45°$. FOBA (dashed curves in Fig. 5.2) predicts a vanishing R at $\theta = \theta_B$, while the exact solution gives a non-zero (but very deep) minimum.

The condition $a \geq \lambda/2q_p$ is fulfilled by the 400 nm-wide defect but it is not fulfilled by the 200 nm-wide defect. Correspondingly, the interference related dip is only observed in the former case. Notice that FOBA only gives an estimation for the interference reflection minimum; for the 400 nm-wide defect, FOBA predicts the dip at $\theta_g = 43.9°$, while the minimum in the full calculation appears at $\theta_g = 37.4°$.

The out-of-plane emittance in Fig. 5.2 shows the same pattern for the two defect widths considered. S initially increases with θ up to a maximum, whereafter it decreases and vanishes at the critical angle of incidence θ_c. For angles larger than θ_c, the wavevector component, k_y (which is conserved in the scattering), is larger than g, so there are no radiative modes the SPP can couple to. Notice that this critical angle appears for all frequencies at which SPP exists, as the SPP has an in-plane wavevector larger than g. Clearly, the closer the SPP dispersion relation is to the light line, the closer θ_c gets to $\pi/2$. The out-of-plane emission can also be expressed in terms of the angle ϕ in Fig. 5.1, which characterizes the sector of possible directions of radiation within the continuum of propagating plane waves. Its dependency upon the angle of incidence and surface impedance, $\phi = 2\arccos(q_p \sin\theta)$, comes from simple geometrical considerations. At the critical angle of incidence, $\theta_c = \arcsin(1/\sqrt{1-Z_s^2})$, ϕ vanishes and, consequently, $S = 0$. In the region where $\theta \geq \theta_c$ only in-plane "elastic" SPP scattering is taking place. For the parameters used in Fig. 5.2, $\theta_c = 73.91°$.

Fig. 5.2 also shows that the difference between the results of GTA and Rayleigh expansion increases with the angle of incidence. At larger θ, the SPP is diverted on larger paths across the metal defect; therefore the influence of the different treatment of boundary conditions by the two methods becomes more evident. In order to stress the difference between impedance and surface relief scatterers, Fig. 5.3 presents GTA calculations for an aluminum protrusion of rectangular shape, located onto a silver surface. This inhomogeneity actually possesses the scattering properties inherent to both impedance and relief defects. No Brewster-like reflectance minima is exhibited by a relief perturbation, since the polarization induced in the defect points mainly along the direction normal to the plane. However, the interference factor $e^{\mathrm{i}k_{px}a} - 1$ also appears

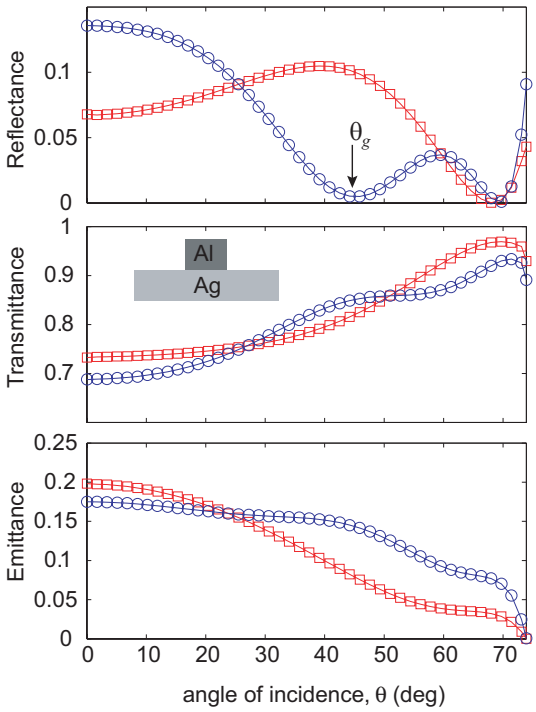

Figure 5.3: (Color online) Scattering coefficients for an air-silver SPP impinging onto an Al protrusion of height 40 nm, as a function of angle of incidence. The squares (circles) are for protrusions of width 200 (400) nm.

in this case in the reflection coefficient of the FOBA, so that a minimum in reflectance is expected at $\theta = \theta_g$. This minimum does in fact appear for the aluminum defect of the 400 nm width, as shown in the numerical results in Fig. 5.3. Also note that the values of the reflectance and emittance for the scattering by a relief defect are much larger than those for the scattering by an impedance defect. This follows strictly from the form of the electric fields produced by dipoles in the metal or vacuum half-spaces. As we have explained, the effective polarization of the impedance defect points parallel to the metal surface, whereas that of the relief defect points perpendicularly to the surface. Therefore, according to eqs.(3.33), (3.34) of Sec.3.4.3 the modulus of the electric field corresponding to the dipole in the metal is $|Z_s|$ times smaller than that corresponding to the dipole in the vacuum. Finally, Fig. 5.4 illustrates the influence of the absorption on the reflection of SPPs by an impedance defect. Reflectivity is calculated using the GTA. As Fig. 5.4 shows, absorption has a small effect on the reflection coefficient, whose amplitude is slightly reduced. The period of oscillations in the defect width is virtually unaltered.

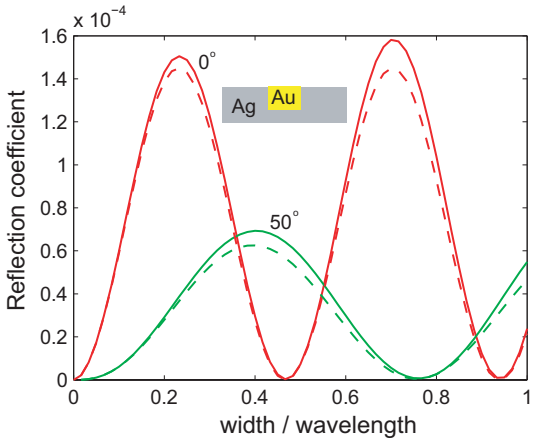

Figure 5.4: (Color online) The SPP reflection coefficient R, as a function of the defect width for two angles of incidence. The SPP impinges onto the gold rectangular inhomogeneity (with the depth 40 nm) along the silver surface. The wavelength is 600 nm. The solid (dashed) curves are for the defect with (without) absorption.

5.3 Conclusions

We have presented a theoretical investigation of oblique SPP scattering by 1D impedance defects and have contrasted the results with those of relief defects. Calculations have been performed within the Green's Dyadic method (which is virtually exact, if small enough values of the mesh are used) and compared with an approximate treatment (developed by A.Yu.Nikitin) where surface impedance boundary conditions have been imposed. The quality of the agreement obtained with both methods validates the theoretical result. The presented results display a net distinction between SPP scattering by impedance defects and SPP scattering by protrusions. While scattering by impedance defects presents a Brewster-like deep reflectance minimum, for angles of incidence $\theta \approx \pi/4$, scattering by protrusions does not show this effect. Such distinction has been shown to arise from the different polarizations induced by the incoming SPP at the position of the defect. The polarization induced in an impedance barrier points in the direction of incidence mainly, while the polarization induced in a protrusions predominantly points perpendicular to the surface. Both protrusions and impedance defects of rectangular shape exhibit low reflectance for certain angles of incidence for which the SPP waves reflected from the two ends of the barrier interfere destructively.

Chapter 6

Comparative study of surface plasmon scattering by shallow ridges and grooves

6.1 Introduction

Surface Plasmon Polaritons (SPP) are electromagnetic bound modes responsible for the transport of light at the interface separating a metal from a dielectric. Their ability to confine light at an air-dielectric interface offers the prospect of developing a new technology consisting of photonic nano-devices[72–75]. Active research is currently focusing on the possibility of achieving control over the propagation of SPPs by means of optical elements that would couple or decouple light to them[56, 69, 76–78]. In order to conceive optical elements (lenses, mirrors, beam-splitters) able to manipulate SPPs propagation, we need to learn more about the interaction of surface plasmons with a sub-wavelength modification of the underling dielectric metal interface. Indeed the interaction of SPPs with surface sub-wavelength defects on a metal surface is of great interest from a theoretical standpoint.[10, 79]
In this chapter we shall study scattering of SPPs by a *shallow* surface defect. We will consider both indentations of the metal surface (grooves) and protrusions on it (ridges). We shall only deal with bi-dimensional defects, which are deemed infinite in one dimension parallel to the interface (the y-direction). Different aspects of this problem have been studied before with a variety of numerical techniques [68, 80–87]. Here we present a systematic comparison between the different scattering coefficients and provide both analytical expressions and qualitative explanations.
It must be noted that a previous work has presented such a comparison[70]

but within an approximate numerical scheme. Within that framework it was found that ridges and grooves exhibited the same scattering, whenever they are shallow enough. Here we will revise that result, which turns out to be valid only for long (elongated) defects. The mistaken outcome of Ref. [70] for short defects may be traced back to the breakdown of the assumption of small curvature in the defect geometry that was made there. In this thesis we solve the Maxwell equations through a discretization method, which does not assume the previous approximation and whose accuracy depends only on the discretization mesh. We found that, as in the previous work, long asymmetric ridges or grooves with the width much larger than the depth, do scatter very similarly. However square shallow defects manifest a different scattering efficiency, differing in the relative radiative loss and radiation pattern. The lack of distinction between these two cases did not emerge in the previous approximate treatment. On the whole the problem needs to be revisited so as to: *i*) substantiate why the approximate result does work in the case of elongated defects, *ii*) point out what is the correct result in the case of shallow and short symmetric defects, and *iii*) explain qualitatively how the scattering properties of short and shallow symmetric defects are gradually transformed into the scattering properties of elongated defects as the aspect ratio of the defect increases.

This chapter is organized as follows. In Sec. 6.2 we state the basic assumptions on the scattering system as well as the solution method. In Sec. 6.3 we rearrange the asymptotic expansions of the far-field to produce the scattering coefficients. Namely we express the far-field and the related Poynting vector in terms of the field inside the defect. Still in this section we look at an approximation for the scattering coefficients of shallow ridges. In Sec. 6.4 we explain that in general we cannot quantitatively represent a scatterer by one mesh (however small). We tell how we associate a small symmetric ridge or groove to a point dipole. In Sec. 6.5 we look at exact numerical results for the scattering of shallow defects of various horizontal lengths. We analyze these results and, in the case of square defects, we associate a ridge to a vertical dipole and a groove to a horizontal dipole. In Sec. 6.6 we produce an analytical model that explains the radiation pattern of the surface plasmons scattered by small square ridges and grooves. In Sec. 6.7 we look at the solutions for the case of shallow and long defects and we present a clear-cut interpretation to support the results of the previous treatment[70]. Finally in Sec. 6.8 we explain qualitatively that the aspect ratio of the defect determines the orientation of the field induced in a shallow defect.

6.2 The scattering systems considered

The considered defects are infinite in the y-dimension and shallow $h << \lambda$. They are going to be illuminated by a monochromatic surface plasmon at normal incidence $\mathbf{e}_{spp}$, associated to an impinging energy flux S_{spp}, defined and derived in Sec.3.3.1. Therefore, only radiation into p-polarized(TM) waves needs to be considered. The material making the slab shall be lossless silver[61],

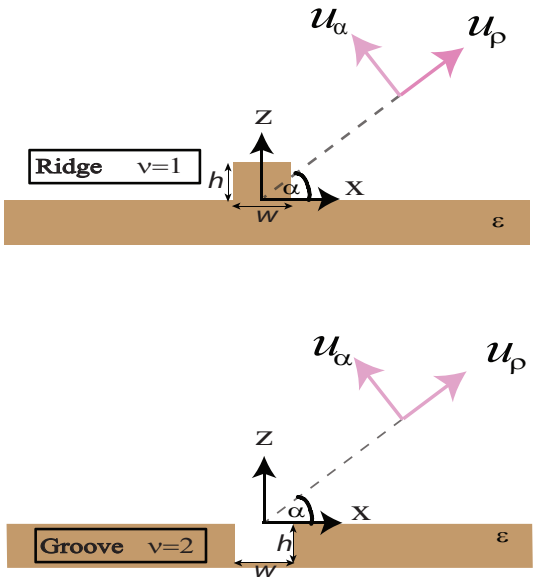

Figure 6.1: Schematic representation of the scattering systems considered. A ridge is a one dimensional defect located in air and labeled by the index $\nu = 1$. A groove is a one dimensional defect located in the metal and labeled by the index $\nu = 2$

that is: $\varepsilon = \Re\{\varepsilon_{Ag}(\lambda)\}$. Absorption is neglected as we consider non-resonant defects with widths much smaller than the SPP propagation length.
As represented in Fig. 7.1, we shall be expressing the source orientation in a cartesian basis $(\mathbf{u}_x, \mathbf{u}_z)$, and the scattered fields in a right-handed orthogonal polar basis:

$$\mathbf{u}_\rho = \cos\alpha\mathbf{u}_x + \sin\alpha\mathbf{u}_z \quad (6.1)$$

$$\mathbf{u}_\alpha = -\sin\alpha\mathbf{u}_x + \cos\alpha\mathbf{u}_z \quad (6.2)$$

Finally a question of notation: throughout we shall refer to a bi-dimensional point-source simply as a dipole, but it is meant that their emission as all of the fields are cyclical in the y-direction. As represented in Fig. 7.1, each object lying in the vacuum semi-space shall be labeled by the superscript $\nu = 1$ while any object lying in the metal shall be labeled by the superscript $\nu = 2$. In particular, scattering quantities related to ridges have the superscript $\nu = 1$

while the ones related to grooves have the superscript $\nu = 2$. The field within the cross-sectional area of the ridge is labeled $\mathbf{E}^r(\boldsymbol{\rho}')$ and the one within that of the groove is labeled $\mathbf{E}^g(\boldsymbol{\rho}')$.

6.3 Scattering Coefficients

In Chapter 3 we have arrived at an explicit expression for the scattered electric far-field. This is attained by propagating the field induced by a dipole density $\mathbf{P}^{(1)}(\boldsymbol{\rho}') = \Delta\varepsilon\ \mathbf{E}^r(\boldsymbol{\rho}')$ (where $\Delta\varepsilon = \varepsilon - 1$) inside the area of a ridge, to a point $\boldsymbol{\rho}$ very far from the source. For a groove we have the same relation between polarization and field (except for a change of sign) $\mathbf{P}^{(2)}(\boldsymbol{\rho}') = -\Delta\varepsilon\ \mathbf{E}^g(\boldsymbol{\rho}')$. To propagate the field from any of the two, we use the standard formula of Eq.(3.4). One of the advantages of the Green tensor technique is that once the fields inside the defects $\mathbf{E}(\boldsymbol{\rho}')$ (and thus $\mathbf{P}(\boldsymbol{\rho}')$) are computed numerically the asymptotic expansions of scattered fields become analytic. This takes us to our first task of this chapter, which is making a direct connection between the orientation of the induced polarization inside the defects and the far-field radiation pattern, and in so doing define the scattering coefficients.
First of all, finding the scattered electric far-field $\mathbf{E}_{rad}(\boldsymbol{\rho})$ requires the asymptotic expansions of the Green tensor. The derivation was given in Sec.3.5, at oblique incidence $k_y \neq 0$. In what follows we give some simplifying rearrangements that will let us focus directly on the angular radiation pattern of surface defects at normal incidence ($k_y = 0$).

6.3.1 Scattering into Radiative Modes

The asymptotic Green's tensor in the radiative zone was derived in Sec.3.5. At normal incidence it can be written in a compact form, for either a ridge or a groove, as:

$$\hat{\mathbf{G}}^{(\nu)}(\rho \to \infty, \alpha, \boldsymbol{\rho}') = \frac{e^{i(g\rho+\pi/4)}}{\sqrt{8\pi g\rho}}\, e^{-igx'\cos\alpha} e^{-ik_z^{(\nu)}z'}\ \hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \boldsymbol{\rho}'). \tag{6.3}$$

In such form we can factor the asymptotic scalar green function out of the dyadic part of the Green tensor. The direction of $\mathbf{E}_{rad}(\boldsymbol{\rho})$ results from superposition of $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^\nu(\boldsymbol{\rho}')$, the emission from all induced point polarization elements, or dipole density elements. Yet the direction of each contribution $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^\nu(\boldsymbol{\rho}')$ must be independent of $\boldsymbol{\rho}'$. In other words since electromagnetic waves are transverse waves in vacuum, far from their source, the field emitted by a dipole $\hat{\mathbf{G}}_\infty^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^\nu(\boldsymbol{\rho}')$ must be proportional to $\mathbf{u}_\alpha$. In fact

setting $k_y = 0$ in the asymptotic expansions of Sec.3.5, we can write:

$$\hat{\mathbf{G}}_{\infty}^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^{\nu}(\boldsymbol{\rho}') = -\left[\boldsymbol{\Pi}^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^{\nu}(\boldsymbol{\rho}')\right] \mathbf{u}_{\alpha}. \tag{6.4}$$

Where for a ridge:

$$\boldsymbol{\Pi}^{(1)}(\alpha, z') = \mathbf{k}_p^{+}(\alpha) + \mathbf{k}_p^{-}(\alpha)\, r_p(\alpha)\, e^{2i\kappa z' \sin\alpha}. \tag{6.5}$$

and for a groove:

$$\boldsymbol{\Pi}^{(2)}(\alpha) = t_p^{(1,2)}(\alpha)\, \mathbf{k}_{pm}(\alpha). \tag{6.6}$$

The expression for the Fresnel coefficients for an air metal interface are given in Eq.(1.121) and Eq.(1.124). In this chapter we only deal with the reflection coefficient for a p-wave propagating from air to metal and for the sake of tidiness, throughout, we omit the superscript of the reflection coefficient $r_p = r_p^{(1,1)}$.

The vectors $\mathbf{k}_p(\alpha)$ are p-waves defined in vacuum, while $\mathbf{k}_p^{m\pm}(\alpha)$ are defined in the metal. Their expressions at normal incidence, in terms of the angle α of Fig. 7.1, particularly, in the far field when $\mathbf{k}/g = \mathbf{u}_r$ in terms of the direct space polar angle α is obtained by noticing that $k_x = g\cos\alpha$ and $k_z = k_z^{(\nu=1)} = g\sin\alpha$ in the air semi-space and $k_z^m = k_z^{(\nu=2)} = g\sqrt{\varepsilon - \cos\alpha^2}$ in the metal. Hence:

$$\begin{aligned}
\mathbf{k}_p^{\pm}(\alpha) &= \frac{1}{g}\left(k_z \mathbf{u}_x \mp k_x \mathbf{u}_z\right) = \sin\alpha \mathbf{u}_x \mp \cos\alpha \mathbf{u}_z \\
\mathbf{k}_p^{m\pm}(\alpha) &= \frac{1}{\sqrt{\varepsilon} g}\left(k_z^m \mathbf{u}_x \mp k_x \mathbf{u}_z\right) = \\
&= \sqrt{\frac{|\varepsilon| + \cos^2\alpha}{|\varepsilon|}} \mathbf{u}_x \mp \frac{\cos\alpha}{\sqrt{\varepsilon}} \mathbf{u}_z
\end{aligned} \tag{6.7}$$

We are now in a position to write the expressions for the radiative fields. Plugging Eq.(6.3) and Eq.(6.4) into Eq.(3.4) we can separate the electric far field dependence into its radial and angular parts as:

$$\mathbf{E}_{rad}^{(\nu)}(\rho, \alpha) = -g^2 \frac{e^{i(g\rho - \pi/4)}}{\sqrt{8\pi g\rho}} E_{rad}^{(\nu)}(\alpha)\, \mathbf{u}_{\alpha} \tag{6.8}$$

here the angular amplitude is:

$$E_{rad}^{(\nu)}(\alpha) = \int_A d\boldsymbol{\rho}'\, e^{-igx'\cos\alpha} e^{-ik_z^{(\nu)} z'} \boldsymbol{\Pi}^{(\nu)}(\alpha, \boldsymbol{\rho}') \cdot \mathbf{P}^{(\nu)}(\boldsymbol{\rho}') \tag{6.9}$$

In the last expression the scattered field in the far zone consists of a cylindrical wave, transverse to the direction of propagation $\mathbf{u}_R$, and with a net angular

amplitude determined by the integral over the source region $E_s^{(\nu)}(\alpha)$. The latter is actually the important bit in the formula as its squared module determines the radiation pattern. As seen from Eq.(6.9) this angular amplitude results from the superposition of each scattering element taken with its own amplitude, phase and optical path in analogy to how an antenna array determines its effective radiation pattern. The radiation is given by the intensity or Poynting vector in the far field. Accordingly the differential angular scattering cross-section is :

$$\frac{\partial \sigma_{rad}^{(\nu)}(\alpha)}{\partial \alpha} = \frac{|\mathbf{E}_{rad}^{(\nu)}(\rho,\alpha)|^2 \rho}{S_{spp}} = \frac{g^3}{S_{spp}} |E_{rad}^{(\nu)}(\alpha)|^2. \tag{6.10}$$

Finally, the net radiative loss σ_{rad} is defined as the integrated angular radiation:

$$\sigma_{rad} = \int_0^{180^0} d\alpha \, \frac{\partial \sigma_{rad}^{(\nu)}(\alpha)}{\partial \alpha}. \tag{6.11}$$

6.3.2 Shallow defects and Green's tensor boundary conditions

Whenever the height of the defect is small enough, typically much smaller than the wavelength of the incident light, we can make the approximation $g|\boldsymbol{\rho}'| << 1$. That allows for some simplification for the angular amplitude of a scattering element *above the surface*. Consider:

$$\begin{aligned} \mathbf{\Pi}^{(1)}(\alpha, \boldsymbol{\rho}') &= (\mathbf{k}_p^+(\alpha) + \mathbf{k}_p^-(\alpha)\, r_p(\alpha)\, e^{2i\kappa z' \sin\alpha}) \simeq \\ &\simeq (\mathbf{k}_p^+(\alpha) + \mathbf{k}_p^-(\alpha))\, r_p(\alpha) \end{aligned} \tag{6.12}$$

Hence, for shallow defects the Green Tensor dependence of Eq.(6.3) is entirely given by the exponential factors $e^{-igx'\cos\alpha} e^{-ik_z^\nu z'}$, for both a source in the vacuum semi-space and a source in the metal semi-space. Indeed this turns out to be a major simplification for the relative amplitude of the scattering elements in the air semi-space, which we shall perform in detail Section6.6.
Before that we need to highlight the relation between the Green tensor of a defect on the metal slab and in the metal slab, under this approximation. Such relation emerges from the boundary conditions for the Green's tensor at the interface, which are:

$$\left[\hat{\mathbf{G}}(\boldsymbol{\rho}, x', z' = 0^+) - \hat{\mathbf{G}}(\boldsymbol{\rho}, x', z' = 0^-)\right] \cdot \mathbf{u}_x = 0 \tag{6.13}$$

$$\left[\hat{\mathbf{G}}(\boldsymbol{\rho}, x', z' = 0^+) - \varepsilon\, \hat{\mathbf{G}}(\boldsymbol{\rho}, x', z = 0^-)\right] \cdot \mathbf{u}_z = 0. \tag{6.14}$$

Notice that, in the unperturbed system, space is translationally invariant in the horizontal direction x and this is reflected in is the x-component of the vector in Eq.(6.12). Because of Eq.(6.3) and Eq.(6.4), we can turn Eq.(6.13) into:

$$\Pi_x^{(1)}(\alpha) = \Pi_x^{(2)}(\alpha) = \Pi_x(\alpha). \tag{6.15}$$

The presence of surface charges at the interface implies, from Eq.(6.14), that the z-components of the vector $\mathbf{\Pi}^{(\nu)}(\alpha)$ on either sides of the interface have the relation:

$$\Pi_z^{(1)}(\alpha) = \varepsilon\, \Pi_z^{(2)}(\alpha). \tag{6.16}$$

6.3.3 Scattering into Surface Plasmons

Let us derive the scattering coefficient into surface plasmon modes. Note that, in this one dimensional problem, scattering will be into both the forward surface plasmon $\mathbf{e}_{spp+}(\boldsymbol{\rho})$, propagating in the positive x direction and the backwards plasmon $\mathbf{e}_{spp-}(\boldsymbol{\rho})$ propagating in the negative x direction, as defined in Sec.3.3.1. The emission by a point dipole or a point polarization element must result into a plasmon final state: $\hat{\mathbf{G}}_{p\pm}(\boldsymbol{\rho}, \boldsymbol{\rho}') \cdot \mathbf{P}(\boldsymbol{\rho}') \propto \mathbf{e}_{spp\pm}(\boldsymbol{\rho})$, as shown in Sec.3.4. The asymptotic Greens tensor for a source *above* ($\nu = 1$) or *in* ($\nu = 2$) the metal is:

$$\hat{\mathbf{G}}_{p\pm}^{(\nu)}(\boldsymbol{\rho}, \boldsymbol{\rho}') \cdot \mathbf{P}^{(\nu)}(\boldsymbol{\rho}') = \frac{-\mathtt{i}}{2g\, S_{spp}} \times \left[\left(\mathbf{e}_{spp\pm}^{(\nu)}(\boldsymbol{\rho}') \right)^* \cdot \mathbf{P}^{(\nu)}(\boldsymbol{\rho}') \right] \, \mathbf{e}_{spp\pm}(\boldsymbol{\rho}).$$

Notice that $\left(\mathbf{e}_{spp\pm}^{(\nu)}(\boldsymbol{\rho}') \right)^*$ complies with Eq.(6.13) and Eq.(6.14). Now, if the the source $\mathbf{P}(\boldsymbol{\rho}')$ is produced by an incident surface plasmon field (as is our case), we can define the scattering cross-section of into SPPs as in Sec.3.4:

$$\sigma_{p\pm}^{(\nu)} = \frac{\mathbf{E}_{p\pm} \times \mathbf{H}_{p\pm}^* \cdot \mathbf{u}_x}{\mathbf{e}_{p\pm} \times \mathbf{h}_{p\pm}^* \cdot \mathbf{u}_x} = \left| E_{p\pm}^{(\nu)} \right|^2 \tag{6.17}$$

where

$$E_{p\pm}^{(\nu)} \;=\; -\frac{\mathtt{i}g}{2\, S_{spp}} \int_{A^{(\nu)}} d\boldsymbol{\rho}' \; \mathbf{e}_{spp\pm}^*(\boldsymbol{\rho}') \cdot \mathbf{P}(\boldsymbol{\rho}') \tag{6.18}$$

Finally, we can define the total scattering cross-section, which in the lossless case is equivalent to the extinction cross-section:

$$\sigma_{xtn} = \sigma_p^+ + \sigma_p^- + \sigma_{rad}. \tag{6.19}$$

6.4 Rayleigh-limit: cautionary remarks

Next we are going to develop solutions to point sources in a metal plane background. However one question may be raised : how do we associate the field induced by a surface plasmon inside a ridge or a groove to a point dipole? The answer is the argument of this section.
When the field inside a defect is obtained by mesh discretization we assume that the field inside a single mesh is uniform, and deviations from the field at its center are deemed negligible. Yet, in general, the field in a defect, cannot be represented by the field at its center alone. Let us explain a little bit further this point.
For simplicity let us consider a defect in a homogenous medium with dielectric constant ε_b, but the argument is the same in other backgrounds. As usual[88], the field at every mesh is found by solving self-consistently a system of N coupled equations :

$$\mathbf{E}(\boldsymbol{\rho}_i) = \mathbf{E}_b(\boldsymbol{\rho}_i) \; + \; g^2 \sum_{j \neq i} \hat{\mathbf{G}}_b(\boldsymbol{\rho}_i - \boldsymbol{\rho}_j) \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_j) \frac{A}{N^2} + \\ + g^2 \, \hat{\mathbf{M}} \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_i) - \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \Delta\varepsilon \, \mathbf{E}(\boldsymbol{\rho}_i) \tag{6.20}$$

where $i = 1, N$ and $j = 1, N$, $\mathbf{E}(\boldsymbol{\rho}_i)$ is the field at the mesh center. $\hat{\mathbf{L}}$ is a term related to the depolarization of light and comes about from the quasi-static contribution of the Green tensor. $\hat{\mathbf{M}}$ is a correction term also coming from static fields.
In practice, the number of mesh points N is increased until the calculation converges to the required precision. Then scale variations $\sim \sqrt{A}/N$ of $\mathbf{E}(\boldsymbol{\rho})$ are properly represented in the solution. In the Rayleigh limit, for a defect of area A so small that $g^2 A << 1$, the scatterer behaves like a point source or a point dipole and the background field (in this case the illumination) can be considered uniform over A: $\mathbf{E}_b(\boldsymbol{\rho}) = \mathbf{E}_b$. Exceptionally, for a circular defect in a homogenous medium with dielectric constant ε_b, the net field at any point $\boldsymbol{\rho}_i$ converges to:

$$\mathbf{E} = \mathbf{E}_b \; - \frac{\hat{\mathbf{L}}}{\varepsilon_b} \cdot \Delta\varepsilon \, \mathbf{E} \tag{6.21}$$

This is because for the field inside an infinitesimal (very sub-wavelength) circular shape is actually uniform and thus scattering by such circular defects can be described by one mesh. In fact the extinction coefficient[31, 38] can be derived from the field at the center alone:

$$\sigma_{xtn} \;\; = \;\; g \Im \left[\int_A d\boldsymbol{\rho}' \, \Delta\varepsilon \mathbf{E}_b^*(\boldsymbol{\rho}') \cdot \mathbf{E}(\boldsymbol{\rho}') \right] = \tag{6.22}$$

$$= \;\; A \Im \left[\Delta\varepsilon \mathbf{E}_b^* \cdot \mathbf{E} \right] \tag{6.23}$$

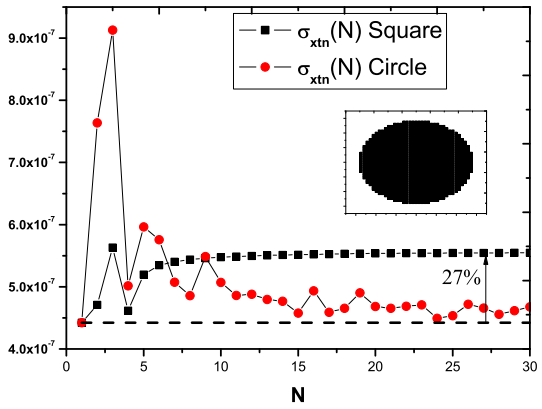

Figure 6.2: Dependence of the extinction coefficient on the number of meshes used in the calculation, for a square defect with $w = h = 1nm$ in vacuum, and a circular defect of the same area, illuminated by a plane wave. The dashed line represents extinction coefficient calculated with the Mie theory for the circle. The dielectric constant in the defect is $\varepsilon = -19.89$ at the wavelength of $700nm$. The inset represents the geometry of a discretized circle when inscribed in a square represented by 30x30 mesh points.

To prove this numerically we have calculated σ_{xtn} for a cylinder represented by a single mesh, as in Eq.(6.21), and illuminated by a plane wave. First of all we have checked that the one-mesh cross-section of Eq.(6.23), coincides with the Mie theory result. Secondly, and we have subdiscretized the cylinder into square meshes as rendered in the inset of Fig. 6.2. As also rendered in the figure, applying Eq.(6.20) we found that, as the number of meshes grows, the scattering cross-section calculated by the collection of meshes Eq.(6.22) converges to the initial value of one single mesh of Eq.(6.23). However the field inside of a square scatterer can never be uniform if it is to satisfy real boundary conditions even in a homogenous medium or vacuum. Thus, it can not be faithfully described by one mesh. This is illustrated in Fig. 6.2, which renders the extinction coefficient for a square defect of the same area as the circle. As it turns out, the converged value is $\sim 27\%$ larger than that obtained by the one-mesh approximation. Remarkably this error is not reduced with the defect size: we obtained the same error for squares with side $5nm$ or $0.5nm$. This is just for reference in the optical range, since we found that the error actually depends on type of defect and on the dielectric constant.
However, even if the field is not uniform, a small defect in the Rayleigh limit can be represented by a point source at the center of the mesh, with its field equal to the average field over the mesh $\overline{\mathbf{E}} = \int_A d\boldsymbol{\rho}' \, \mathbf{E}(\boldsymbol{\rho}')$.

Indeed if $\mathbf{E}_b(\boldsymbol{\rho})$ is negligible over the area of the defect we have:

$$\sigma_{xtn} = A\Im\left[\Delta\varepsilon\mathbf{E}_b^* \cdot \overline{\mathbf{E}}\right] \tag{6.24}$$

So the object behaves as a point-dipole $\mathbf{p} = A\Delta\varepsilon\overline{\mathbf{E}}$.
The previous results were for a homogeneous background, but they also hold for the inhomogeneous one considered in this chapter. We find that, for a defect above the surface in the optical range, the relative error is about 40%, while it can reach 50% for a defects below the surface.
With very small non-elongated ridges and grooves, such that $w \approx h/\lambda << 1$, the equivalent point dipoles are attained by averaging the fields over the area of the defects as follows:

$$\mathbf{p}^{(1)} = \Delta\varepsilon\overline{\mathbf{E}}^r A = \Delta\varepsilon \int_A d\boldsymbol{\rho}' \, \mathbf{E}^r(\boldsymbol{\rho}') \tag{6.25}$$

$$\mathbf{p}^{(2)} = -\Delta\varepsilon\overline{\mathbf{E}}^g A = -\Delta\varepsilon \int_A d\boldsymbol{\rho}' \, \mathbf{E}^g(\boldsymbol{\rho}') e^{-ig|z'|\sqrt{|\varepsilon|}}. \tag{6.26}$$

Accordingly if we set $\mathbf{P}^{(\nu)}(\boldsymbol{\rho}) = \delta(\boldsymbol{\rho} - \boldsymbol{\rho}')\mathbf{p}^{(\nu)}$ Eq.(6.9) and Eq.(6.18) for small non-elongated defects become:

$$E_{rad}^{(\nu)}(\alpha) = \mathbf{\Pi}^{(\nu)}(\alpha) \cdot \mathbf{p}^{(\nu)} \tag{6.27}$$

$$E_{p\pm}^{(\nu)}(\alpha) = \left[\mathbf{e}_{p\pm}^{(\nu)}(0)\right]^* \cdot \mathbf{p}^{(\nu)} \tag{6.28}$$

6.5 Numerical results

As an illustration consider a square ridge and a groove of side $w = h = 10nm$. We have calculated the scattering into radiative modes and SPPs without associating the defect to a point dipole but rather using Eq.(6.10) and Eq.(6.17). Similar numerical results for the case of shallow grooves were found in Ref. [86] using a different computational technique.
The out of plane radiation pattern of a surface plasmon scattered by such defects is given in Fig. 6.3.
Calculations show that, for symmetric defects, the net radiative loss is greater for a groove than for a ridge. This is so because, while both the scattering into SPPs and the radiation close to the surface (at $\alpha = 0, 180°$) are similar, their radiation patterns greatly differ normal to the surface ($\alpha = 90°$), where the groove radiation is maximum while the ridge radiation goes to zero.

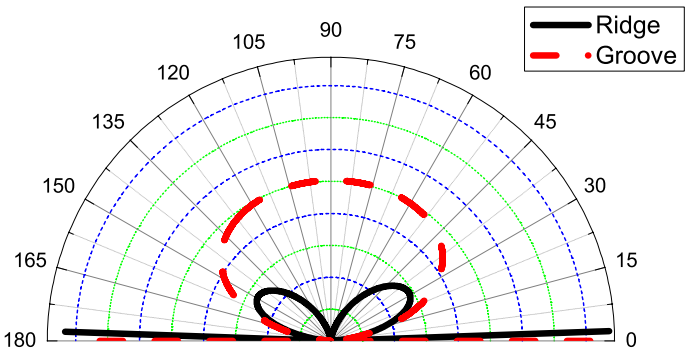

Figure 6.3: Angular radiative cross section (Eq.(6.10)) and surface plasmon cross section (represented by the almost horizontal lines at $\alpha \simeq$ and $\alpha \simeq 180^0$), for square defects with 10nm side, illuminated by a SPP on silver at $500nm$. The scale is linear but the units are arbitrary.

The ridge radiation pattern is distributed into two lobes on either sides of $\alpha = 90°$ but the groove radiation pattern forms a single lobe. Notice the fraction of energy scattered into SPPs is large. This is one of our main result and shall be analyzed in detail in the next section.
These results are not in agreement with those obtained in the approximate treatment Ref.[70]. We associate the discrepancy to the breakdown of the condition that the curvature of a short and shallow defect does not vary rapidly, used in that work.
Let us now keep the defects height at $h = 10nm$ and enlarge the width w. Fig. 6.4 renders the radiation pattern for a rectangular defect of width 50nm ($h = 10nm$). The emergence of directivity in the out of plane radiation, is part of a transitional behavior, in which the radiation patterns tend to align

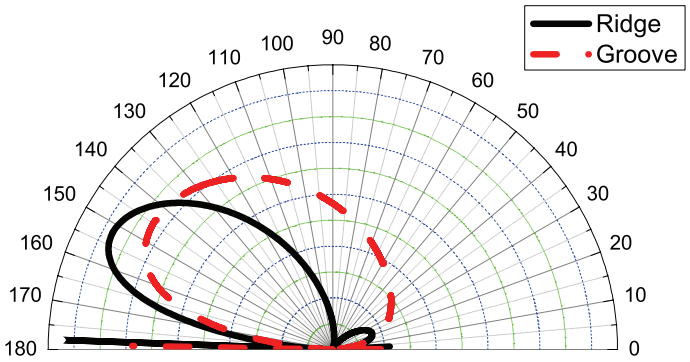

Figure 6.4: Angular radiative cross section (Eq.(6.10)) and surface plasmon cross section (represented by the almost horizontal lines at $\alpha \simeq 0$ and $\alpha \simeq 180^0$) for rectangular defects with $10nm$ height and 50nm width, illuminated by a SPP on silver at $700nm$. The scale is linear but the units are arbitrary.

and, simultaneously, one of the lobes is shrunk while the other is blown up in the ridge radiation. Notice that the scattered energy into SPPs exhibits the same directivity, going mainly in reflection. Eventually, if we keep enlarging the defects until they are considerably asymmetric the radiation patterns for both ridges and grooves tend to be single overlapping lobes (see Fig. 6.5). Noticeably, the scattering into SPPs is greatly reduced. Such similarity is explainable in the approximate framework presented in Ref.[70] which turns out to be quite acceptable in this limit of large enough defects, as we shall substantiated in Sec. 6.7. In Sec. 6.8 we shall account qualitatively for the transition observed in Fig. 6.4, explaining why the radiation pattern changes when the defects are enlarged.

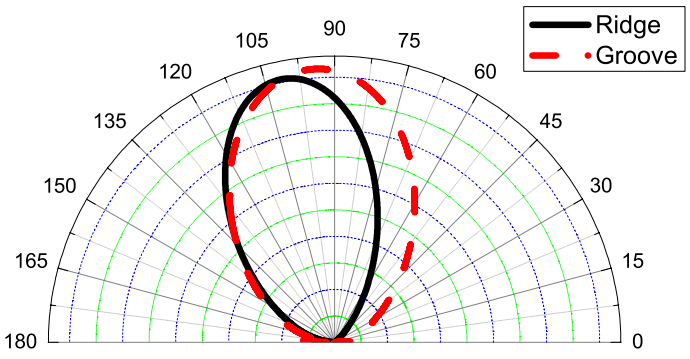

Figure 6.5: Angular radiative cross section (Eq.(6.10)) and surface plasmon cross section (represented by the almost horizontal lines at $\alpha \simeq 0$ and $\alpha \simeq 180^0$), for rectangular defects with $10nm$ height and $300nm$ width, illuminated by a SPP on silver at $700nm$. The scale is linear but the units are arbitrary.

6.5.1 Scattering by square ridges and grooves in the Rayleigh limit

The equivalence between non-elongated subwavelength defects and point dipoles gives us a chance to investigate in depth the individual radiation pattern of a single scattering element.
Fig. 6.6 shows the averaged the field inside the 10nm ridges and grooves, as prescribed in Eq.(6.25) and Eq.(6.26). The field induced in a groove is mainly longitudinal while the field inside the ridge is mainly transversal. This is due to both the illumination and the polarizabilty of the scatterers. When defects are almost symmetric their polarizibilities β_i are nearly isotropic and so the induced field and the incident field are virtually parallel. Hence the field induced in a ridge and a groove are nearly parallel to the incident surface plasmon $\mathbf{e}_{spp}$, which is mainly perpendicular to the plane in the vacuum semi-space and is mainly parallel to the plane in the metal semi-space. Therefore, in the Rayleigh limit, a ridge scatters SPPs into radiative modes like a vertical dipole on the plane, while the groove scatters them into radiative modes like a horizontal dipole on a plane. The results for grooves is in agreement with Ref.[87].
Interestingly, we also have found numerically in Fig. 6.6 that:

$$|\overline{E}_x^g| \sim |\sqrt{\varepsilon}\ \ \overline{E}_z^r| \tag{6.29}$$

especially at short wavelengths. We have devised a virtual source, that can condense the orientation of the equivalent dipole representing a non-elongated symmetric ridges and grooves. This virtual dipole is defined as: $\mathbf{q}(\theta) = \mathbf{u}_x\sqrt{|\varepsilon|}\cos\theta + \mathbf{u}_z\sin\theta$. The fields inside a groove and a ridge, are respectively, represented as:

$$|\overline{\mathbf{E}}^g| \ \simeq\ |\Delta\varepsilon\,\overline{E}_z^r\mathbf{q}(0)| \tag{6.30}$$

$$|\overline{\mathbf{E}}^r| \ \simeq\ |\Delta\varepsilon\,\overline{E}_z^r\mathbf{q}(90^0)| \tag{6.31}$$

at least as long as Eq.(6.29) holds.
In reality we can see what happens by means of Eq.(6.21). Despite the fact that this equation is only exact for a circle in a homogenous background (as explained) we can use it to show *qualitatively* the relation between the field inside the groove and the ridge, when their shapes are symmetric. If we approximate the polarizability of a ridge for that of a circle in vacuum (whose polarizability is calculated through Eq.(6.21)), so $\beta_1 = 2/(\varepsilon+1)$. If we also approximate the groove polarizability by that of a hole in a homogenous metal medium, we have: $\beta_2 = 2\varepsilon/(\varepsilon+1)$. Hence the field induced inside each object

is:

$$\overline{\mathbf{E}}^r \approx \beta_1 \mathbf{e}_{spp}(x=0, z=0^+) = \beta_1 \mathbf{u}_z \tag{6.32}$$

$$\overline{\mathbf{E}}^g \approx \beta_2 \mathbf{e}_{spp}(x=0, z=0^-) = \beta_2 \frac{\mathbf{u}_x}{\sqrt{\varepsilon}} \tag{6.33}$$

Since these polarizabilities also have the property: $\beta_2 = \varepsilon\beta_1$ (the polarizability of a hole in a material is ε times larger than the polarizability of a particle of the same material and the same shape) then $|\sqrt{\varepsilon}\ \overline{E}^g_x| \sim |\overline{E}^r_z|$.

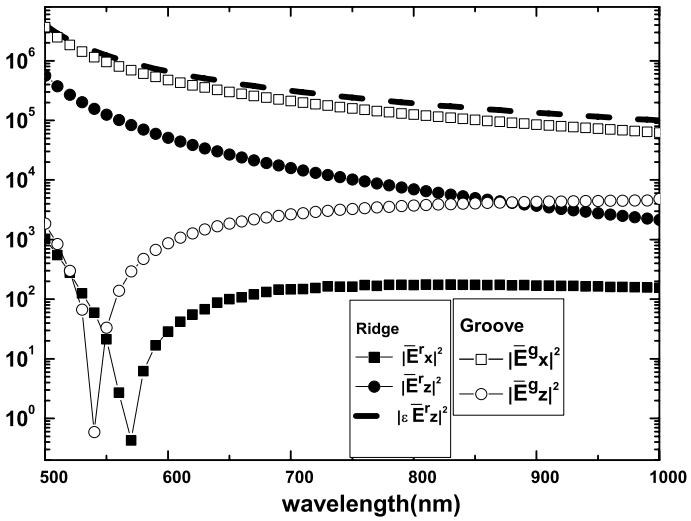

Figure 6.6: The averaged field components (as defined in Eq.(6.25), Eq.(6.26)) for a square groove and a square ridge of 10nm side in silver, as a function of the wavelength. The scale is logarithmic with arbitrary units.

The symmetry of the polarizations β_i and the property $\beta_2 \simeq \varepsilon\beta_1$ are strictly true for circular defects in homogeneous media. Our numerical calculations of Fig. 6.6 shows that, even though the field inside a ridge and a groove are quantitatively different from those of circular defects in homogenous media, the assumption that their mutual relation is preserved is in very good agreement with the exact result. Because of the symmetry of the square shape, the *averaged field* inside the square is very nearly parallel to the incident field.

6.5.2 Reflection of surface plasmons square shallow defects

As a corollary of the properties of the fields in a ridge and a groove $|\sqrt{\varepsilon}\overline{E}^r_z| \sim |\overline{E}^g_x|$ we can also substantiate that their reflection of surface plasmons is quite similar. In fact, we obtain:

$$|E^{(1)}_{p\pm}| \simeq |\overline{E}^r_z| \tag{6.34}$$

$$|E_{p\pm}^{(2)}| \simeq \left| \frac{\overline{E}_x^g}{\sqrt{\varepsilon}} \right| \simeq |\overline{E}_z^r| \simeq |E_{p\pm}^{(1)}| \tag{6.35}$$

Notice that these define $\sigma_p^{\pm}$ through Eq.(6.17). Once σ_{rad} from Eq.(6.11) and σ_p^- are determined the value of the transmission of the surface plasmon is a constrained variable: $T = 1 - \sigma_p^- - \sigma_{rad}$, at least for the lossless case[70]. Since σ_{rad} is greater for grooves than for ridges, the groove transmission is smaller.

6.6 Radiation patterns for Horizontal and Vertical point dipoles on a real metal interface

The first part of the expression Eq.(6.10) is a pre-factor g^3/S_{spp} whereas the second part is the *the radiation pattern* of a point dipole:

$$|E_{rad}^{(\nu)}|^2 = \left| \mathbf{\Pi}^{(\nu)}(\alpha) \cdot \mathbf{p} \right|^2 \tag{6.36}$$

A groove emits like a horizontal dipole. The angular amplitude of the field radiated by a horizontal unit dipole $\mathbf{p} = \mathbf{u}_x$, placed close to the interface $z = 0$, is $\Pi_x(\alpha)$, and it does not matter on which side of the interface it is placed. $\Pi_x(\alpha)$ can be derived using these relations:

$$t_p^{(1,2)}(\alpha) = \frac{2\sqrt{\varepsilon}\sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha} \tag{6.37}$$

$$1 + r_p(\alpha) = \frac{2\sqrt{\varepsilon - \cos^2\alpha}}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha} \tag{6.38}$$

$$1 - r_p(\alpha) = \frac{2\varepsilon\sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha}. \tag{6.39}$$

The explicit result is:

$$\Pi_x(\alpha) = \frac{2\sqrt{\varepsilon - \cos^2\alpha}\ \sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha} \tag{6.40}$$

and the radiation pattern is $|\Pi_x(\alpha)|^2$. Notice $\Pi_x(\alpha)$ presents a mirror symmetry about the angle $\alpha = 90^0$, the normal to to the plane. Furthermore since $\Pi_x(\alpha)$ never changes sign between 0 and 180^0 (nor goes to zero), the field of a horizontal dipole has one single symmetric lobe, where the field always has the same sign. The field intensity $|\Pi_x(\alpha)|^2$ of such lobe is rendered in Fig. 6.7 for different dielectric constants. This radiation pattern of a groove shown in

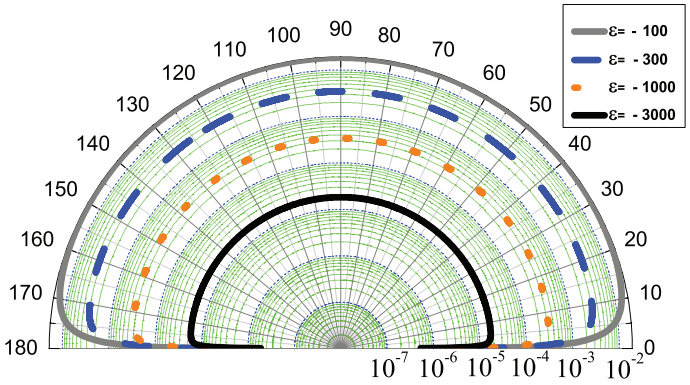

Figure 6.7: Radiative angular intensity $|\Pi_x|^2$ of a horizontal point dipole at an air-metal interface. The radiation patterns varies as the metal dielectric constant is varied. The scale is logarithmic with arbitrary units.

Fig. 6.7, is in agreement with the one represented by Ref.[86], obtained with a different numerical method. Notice that for $|\varepsilon| >> 1$:

$$\Pi_x \to 2\,\varepsilon^{-1/2}. \tag{6.41}$$

This is when ε increases this radiation pattern tends to become simultaneously isotropic and vanishing. In fact a horizontal dipole does not radiate on a perfect conductor[19]. On a small digression it is interesting to notice an apparent contradiction between treatments such as Ref.[89], which considered that a defect in a perfect metal were equivalent to a magnetic dipole, while another work[87] explains a defect in a real metal corresponds to an electric dipole. Actually we have just reconciled the two results. We know that a horizontal dipole on a plane tends to emit isotropically for large ε. This means that on a first order expansion in $1/\varepsilon$, the radiation pattern of a horizontal dipole on a plane and that of a magnetic dipole in vacuum, are identical.
For finite ε the field $\Pi_x(\alpha)$ of a horizontal dipole within a real metal would not be thoroughly screened, and while the pattern remains symmetric, its isotropy is disrupted parallel to the surface (i.e. $\alpha = 0, 180^0$) to accommodate the emergence of the surface plasmons density of states.
For an individual vertical dipole $\mathbf{p} = \mathbf{u}_z$, which represents a ridge, the angular amplitude of the field is:

$$\Pi_z^{(1)}(\alpha) = \frac{2|\varepsilon|\sin\alpha}{\sqrt{\varepsilon - \cos^2\alpha} + \varepsilon\sin\alpha}\quad \cos\alpha \tag{6.42}$$

The field from a vertical dipole also goes to zero at $\alpha = 0, 180^0$ for a finite ε, but since dipoles only radiate transversally, the field has a third zero at 90^0. The field is antisymmetric with respect to the normal of the plane, while the intensity $|\Pi_z(\alpha)|^2$ is symmetric, and is made up of the two lobes separated by

a zero at 90^0, see Fig. 6.8. Yet it is important to keep in mind that the field of one lobe is in anti-phase with the field of the other.

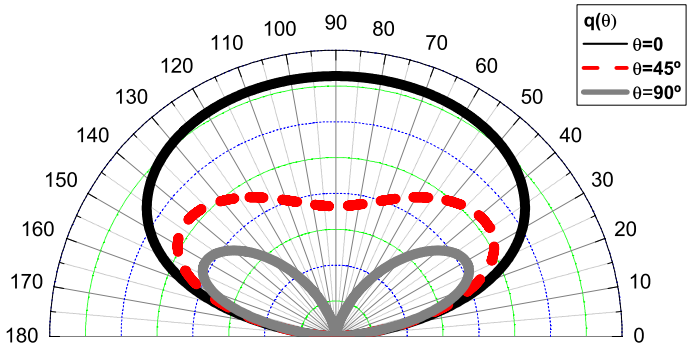

Figure 6.8: Radiation pattern $|E_{rad}(\alpha)|$ from the virtual dipole $q(\theta)$ at $\theta = 0, 45^0, 90^0$. The scale is linear with arbitrary units.

Unlike a horizontal dipole, the vertical dipole radiative field does not vanish for $|\varepsilon| >> 1$ in fact:

$$\Pi_z^{(1)} \to 2\cos\alpha \tag{6.43}$$

The total radiation from a vertical dipole has a larger weight than the radiation by a horizontal one, by a factor of $\sqrt{\varepsilon}$. This can be seen, in fact, from Eq.(6.40) if we assume ε is large, we get the following relation:

$$\Pi_z^{(1)}(\alpha) \simeq \sqrt{\varepsilon}\ \Pi_x(\alpha)\cos\alpha \tag{6.44}$$

In Fig. 6.8 we represent radiation pattern of $\mathbf{q}(\theta)$ for the horizontal and vertical orientations respectively, $\theta = 0$, $\theta = 90^0$, which corresponds to our analytic analog of the emission pattern of square ridges and grooves respectively. While we will consider an intermediate orientation in the next section, we want to remark here that, due to Eq.(6.44), the radiation by both the horizontal moment $\mathbf{q}(0)$ and a vertical moment $\mathbf{q}(90^0)$ vanish parallel to the plane at $\alpha = 0, 180^0$ in a similar manner, as illustrated in Fig. 6.3

At the same time the far-field emissions of ridges and grooves become increasingly different as we approach the direction normal to the plane.

6.7 Solutions for long and shallow Ridges and Grooves

For shallow and long defects $w > h$ and $h/\lambda << 1$ we define the following height-averaged polarization densities and fields:

$$\begin{aligned} \widetilde{\mathbf{P}}^{(1)}(x') &= \Delta\varepsilon \int_0^h dz' \mathbf{E}^{(1)}(x', z') = \\ &= \Delta\varepsilon\, \widetilde{\mathbf{E}}^{(1)}(x')\, h \end{aligned} \tag{6.45}$$

where the last equation defines $\widetilde{\mathbf{E}}^{(1)}(x')$. Likewise for a groove we can define $\widetilde{\mathbf{P}}^{(2)}(x')$ and $\widetilde{\mathbf{E}}^{(2)}(x')$ through the following equation :

$$\begin{aligned} \widetilde{\mathbf{P}}^{(2)}(x') &= -\Delta\varepsilon \int_{-h}^0 dz' \mathbf{E}^{(2)}(x', z')\, e^{-g|z'|\sqrt{|\varepsilon|}} \\ &= -\Delta\varepsilon\, \widetilde{\mathbf{E}}^{(2)}(x')\, h \end{aligned} \tag{6.46}$$

Notice for $|\varepsilon| >> 1$ we can make the approximation $k_{pz}^m \sim k_z^m \sim ig\sqrt{\varepsilon}$.
The benefit of using $\widetilde{\mathbf{P}}^{(\nu)}(x')$ is that the scattered-field coefficients for these defects in the far zone, $E_s^{(\nu)}(\alpha)$ and $E_{p\pm}^{(\nu)}$, are those emitted by a chain of point-dipoles on the surface over the segment w, and set at 0^+ and 0^- for ridges and grooves, respectively.
The scattered field angular amplitude $E_s^{(\nu)}(\alpha)$ from Eq.(6.9) and Eq.(6.12) is obtained as:

$$E_{rad}^{(\nu)}(\alpha) \simeq \mathbf{\Pi}^{(\nu)}(\alpha) \cdot \int_0^w d\boldsymbol{\rho}'\, \widetilde{\mathbf{P}}^{(\nu)}(x')\, e^{-igx' \cos\alpha} \tag{6.47}$$

This holds for the scattering into surface plasmon modes as well since we have:

$$E_{p\pm}^{(\nu)} = \left[\mathbf{e}_{spp\pm}^{(\nu)}(0)\right]^* \cdot \int_0^w dx'\, e^{\mp ik_{px}x'}\, \widetilde{\mathbf{P}}^{(\nu)}(x'). \tag{6.48}$$

When we illuminate a shallow and long defect, with a SPP, an equivalent linear density of dipole sources $\widetilde{\mathbf{P}}(x')$ stems from how the induced fields are distorted inside the scatterer, namely by its polarizability. When the defect is larger in the horizontal direction than in the vertical one, ridges and grooves were found to give the same scattering by an approximated Rayleigh expansion[70]. We have an alternative first principles argument to justify the Rayleigh expansion result, which is based entirely on the assumption that these defects are needle shaped. The field induced in these defects tends to be that induced in a needle-shaped protrusion placed horizontally on the surface 0^+ in the case of a ridge. For a groove we have a horizontal needle-shaped cavity at 0^-. In such idealistic simplification it is clear-cut to deduce the fields inside the defects

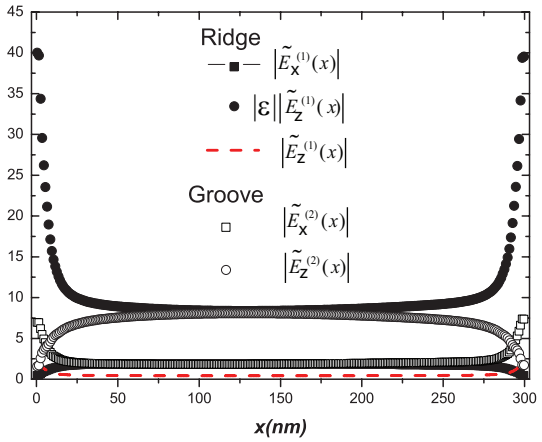

Figure 6.9: The averaged fields component inside of a ridge and a groove, $\widetilde{E}_x^{(1)}$, the quantity $\varepsilon\widetilde{E}_z^{(1)}$ and the rest of the components $\widetilde{E}_x^{(1)}$, $\widetilde{E}_x^{(2)}$, $\widetilde{E}_z^{(2)}$, for rectangular defects of $w = 300nm$ and $h = 10nm$. The system is illuminated by a SPP in lossless silver at $\lambda = 700$nm. The scale is linear with arbitrary units.

from the boundary conditions. Namely the parallel component of the incident field is always continuous and equal, as in Eq.(3.16)and Eq.(3.17):

$$\widetilde{E}_{1x}(x') = \mathbf{e}_{sp}(x', 0) \cdot \mathbf{u}_x = \widetilde{E}_{2x}(x') \tag{6.49}$$

which preserve the continuity of Eq.(6.13). However, we are generating fields which, normal to the surface, make up for the discontinuity perpendicular to the metal surface of Eq.(6.14). In fact, for a horizontal needle-like ridge, the boundary conditions imposed by the continuity of the displacement vector are:

$$\widetilde{E}_{1z}(x') = \mathbf{e}_{sp}(x', 0^+) \cdot \mathbf{u}_z/\varepsilon = 1/\varepsilon \tag{6.50}$$

while for a needle-like slit:

$$\widetilde{E}_{2z}(x') = \varepsilon\, \mathbf{e}_{sp}(x', 0^-) \cdot \mathbf{u}_z = 1. \tag{6.51}$$

Ultimately:

$$\widetilde{E}_{1x}(x') = \widetilde{E}_{2x}(x') \tag{6.52}$$

$$\varepsilon\widetilde{E}_{1z}(x') = \widetilde{E}_{2z}(x') \tag{6.53}$$

which, matched with Eq.(6.13) and Eq.(6.14), yields:

$$\left|\hat{\mathbf{G}}(\boldsymbol{\rho}, x', z' = 0^+) \cdot \widetilde{\mathbf{E}}_1(x')\right| = \tag{6.54}$$
$$= \left|\hat{\mathbf{G}}(\boldsymbol{\rho}, x', z' = 0^-) \cdot \widetilde{\mathbf{E}}_2(x')\right|$$

and thus the property of producing the same scattering coefficients, previously found in Ref.[70]. Of course this is just an approximation, but it explains why

elongated defects have similar scattering properties. In real life the plasmon scattering by protrusions and indentations is similar because, far from the edges, a shallow but elongated defect behaves as an infinitely elongated one, as confirmed by numerical calculations. As an example we report in Fig. 6.9 a numerical calculation of the fields averaged over the height for defects of $w = 300nm$ and $h = 10nm$. This shows that Eq.(6.52) and Eq.(6.53) are quite accurate at the center of the defect, and deviates from the needle model prediction due to fringe effects at the edges.
It is worth mentioning that this equivalence is valid in the Rayleigh limit when the defect size is much smaller than the wavelength, and may be altered at resonant wavelengths.

6.8 The transition from short and shallow defects to long and shallow defects: oblique dipoles on a real metal plane

Everything we just said for symmetric surface defects was based on the fact that their aspect ratio equals one. As the defect width is increased, the aspect ratio becomes larger and this leads, progressively, to an asymmetric polarizability tensor. The first effect is that the field induced is gradually less and less parallel to the incident field. Therefore a ridge would develop a non-negligible horizontal electric field component, thus ceasing to be equivalent to a vertical dipole. Likewise the groove, which in the symmetric case behaves as a horizontal dipole, gradually starts having a non-negligible vertical component as its shape is elongated. The process goes on until we recover the case of a needle shaped defect of section 6.7. The fields inside a defect having intermediate width, as in Fig. 6.4, are intermediate between those for the needle case and the square symmetric case. Therefore in these cases defects emit qualitatively like *oblique dipoles*, with the orthogonal components out of phase.
In order to understand better the radiation pattern by ridges and grooves we decompose the oblique dipole in its horizontal and vertical components.
First of all, we focus on the mechanisms involved radiation pattern for a ridge $\nu = 1$. From Eq.(6.27) a dipole with arbitrary orientation emits close to the surface, with a field angular amplitude:

$$E^{(1)}_{rad}(\alpha) = \mathbf{\Pi}^{(1)}(\alpha) \cdot \mathbf{p}^{(1)} \quad = \quad \Pi_x(\alpha)\Delta^{(1)}(\alpha) \tag{6.55}$$

where $\Delta^{(1)}(\alpha) = p^{(1)}_x + \left(\Pi^{(1)}_z(\alpha)/\Pi_x(\alpha)\right) p^{(1)}_z$ and equals:

$$\Delta^{(1)}(\alpha) = p^{(1)}_x + i\,\frac{\varepsilon\cos\alpha}{\sqrt{\cos^2\alpha + |\varepsilon|}}\,p^{(1)}_z \tag{6.56}$$

$\Delta^{(1)}(\alpha)$ shows that the contribution to the radiative field coming from the vertical and horizontal dipole on a metal plane have a phase difference of 90^0. This was already evident from Eq.(6.44), when $\varepsilon < 0$. Such phase difference arises from the impedance of a metal plane[70] $Z_s = -i/\sqrt{|\varepsilon|}$.
The radiation pattern for a dipole with arbitrary orientation and lying in the top of the metal, is written in our formalism as: $|\Pi_x(\alpha)\Delta^{(1)}(\alpha)|^2$. The net angular amplitude for an oblique dipole is resolved into the superposition of the angular envelope of the horizontal dipole (shown in Fig. 6.7), with the other radiation factor $|\Delta^{(1)}(\alpha)|^2$. This last factor contains both the orientation and phase of the field. To envisage how these combine we may develop $|\Delta^{(1)}(\alpha)|^2$ into three terms. These consist in the individual emission from the horizontal

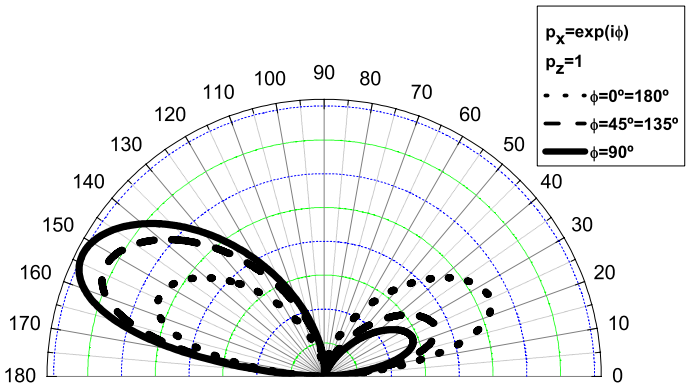

Figure 6.10: Radiation pattern $|E_{rad}(\alpha)|$ for a point dipole: $\mathbf{p} = \mathbf{u}_x e^{i\phi} + \mathbf{u}_z$, lying on top of a metal surface. The scale is linear with arbitrary units.

and vertical dipole plus an interference term:

$$|\Delta^{(1)}(\alpha)|^2 = |p_x|^2 + \frac{|\varepsilon|^2 \cos^2\alpha}{|\varepsilon| + \cos^2\alpha}|p_z|^2 + \\ -2\,\frac{|\varepsilon|}{\sqrt{\cos^2\alpha + |\varepsilon|}}\,\Im\left[p_{1x}p_z^*\right]\cos\alpha \tag{6.57}$$

In the presence of the plane metal background, we have that horizontal and vertical dipoles behave as individual sources but their interaction presents an intrinsic added phase difference of 90^0, which is due to the different interaction of a horizontal and a vertical dipole with the plane. As a result, when in phase they do not interfere, and their radiation pattern is always symmetric regardless of the orientation of the dipole. This is the case for $\mathbf{q}(45^0)$ where, as in Fig. 6.8, the radiation pattern is the sum of the angular intensity of a vertical and a horizontal dipole, so that at 90^0 you have a minimum due to the vanishing of the vertical dipole contribution and yet never goes to zero because of the horizontal dipole contribution. Nevertheless, when the dipole components are not in phase, we can get asymmetric radiation patterns and additional zeros (to those at 0^0 and 180^0), because the interaction term can be negative. In such case the interaction of the horizontal radiative field (with only one lobe) with the vertical radiative (with two lobes of different sign) is responsible for an asymmetric radiation pattern and exhibits directionality. This is illustrated in Fig. 6.10 for a dipole emission whose main contribution comes from the vertical dipole. In Fig. 6.11 we show the radiation pattern for a dipole whose main contribution comes from the horizontal dipole radiation. For the case of a grooves ($\nu = 2$), the radiative angular field amplitude is, from

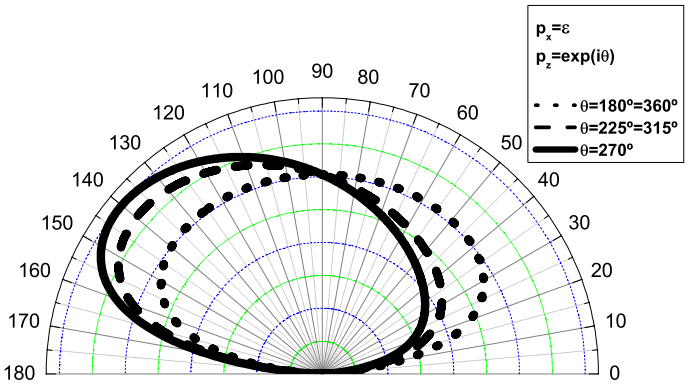

Figure 6.11: Radiation pattern $|E_{rad}(\alpha)|$ for a point dipole: $\mathbf{p} = \varepsilon \mathbf{u}_x + e^{i\phi}\mathbf{u}_z$, lying on top of a metal surface. The scale is linear with arbitrary units.

Eq.(6.16):

$$\mathbf{E}^{(2)}_{rad}(\alpha) = \Pi_x(\alpha)\Delta^{(2)}(\alpha)\, e^{-g|z'|\sqrt{|\varepsilon|}} \tag{6.58}$$

$$\Delta^{(2)}(\alpha) = p^{(2)}_x + i\,\frac{\cos\alpha}{\sqrt{\cos^2\alpha + |\varepsilon|}}\, p^{(2)}_z \tag{6.59}$$

where remember we have also added the approximation: $k^m_z \simeq g\sqrt{\varepsilon}$ for $|\varepsilon| >> 1$.

Remarkably, as opposed to the the dipole emission over the surface, in the net emission from a dipole under the surface the horizontal dipole contribution has a greater weight than the vertical dipole contribution. Apart from this, all the arguments used for a dipole over the surface apply.

The interaction between the vertical and horizontal components of the field induced in the field generates the directional patterns of Fig. 6.4. For a ridge with length slightly larger than its height the directional radiation is dominated by its vertical component. Fig. 6.10 exemplifies the effect of the interference of a dominant vertical component with a smaller but non-negligible horizontal component. For even larger aspect ratios the contribution from the other component may be comparable.

Likewise when a groove has a small aspect ratio it is predominantly a horizontal source interfering with a smaller vertical source. The result is in an interference pattern that looks like the one rendered in Fig. 6.11. Yet again this can be modified by increasing the aspect ratio. This transition is in good agreement with Fig. 11 of Ref.[86] where, using a different numerical method, the radiation pattern of a groove was computed for different aspect ratios.

6.9 Conclusions

Our analysis of the surface plasmon scattering by square shallow defects into radiative modes and plasmon modes, reveals that a groove scatters more of the incident energy than a ridge does. The reflection by a symmetric ridge and a groove is similar and so is the radiative emission close to the horizontal direction. Indeed their scattering essentially differs in the vertical direction, where a groove scatterers while a ridge does not. When defects start to become longer in width we saw the polarizability gets more asymmetric. Correspondingly, since both components of the incident plasmon are out of phase, defects are equivalent to interfering horizontal and vertical dipoles on a plane, which interfere constructively in some direction, thus producing directionality in the radiation pattern. Finally when ridges and grooves are shallow and long they tend to produce the same scattering as, apart for fringe effects, their polarizability exactly counterbalances the discontinuity of the incident surface plasmon field at the air-metal interface.

Chapter 7

Surface plasmon scattering by shallow and deep surface defects

7.1 Introduction

Surface Plasmon Polaritons (SPPs) are bound modes that confine light at the interface separating a metal from a dielectric. The study of surface plasmons is an active research field sometimes referred to as plasmonics[74, 79]. One of the aims of plasmonics is to control the propagation of surface plasmons by means of optical elements that could couple or decouple light to surface plasmons[52, 56, 69, 77, 78], with the prospect of developing a new technology consisting of photonic nano-devices[72, 73, 75].
In this chapter, we present a comparative study of scattering of surface plasmon polaritons by defects of different shapes. The defects can be either indentations of the metal surface (grooves) or protrusions on it (ridges). Some properties of ridges and grooves have been investigated in several works[68, 84, 85, 90–92]. Nonetheless, a clear general picture of what configurations are best for different optical functionalities, has not yet emerged. This is because plasmonic systems, such as ridges and grooves, involve a large number of physical quantities at different scales. As a consequence, dramatic changes may result from slight variations of any of the many parameters involved: wavelength, sizes, shapes, materials, the period in arrays of defects, types of illumination and so forth.
We have presented a comparative study of the scattering of *shallow* ridges and grooves in Chapter 6, a case for which analytical expressions can be obtained. In this chapter we extend that study to consider the height dependence of the scatterers, and focus on analyzing systematic changes on the scattering properties, rather than on the optimization of physical properties. We analyze the scattering of a SPP by both individual defects and arrays of defects, to

elucidate how band-gaps effects affect the properties of individual defects. In order to reveal more clearly the differences between scattering properties of ridges and grooves, in this work we consider the simplest case of bi-dimensional defects, which are deemed infinite in one of the dimensions parallel to the interface.

7.2 Scattering Systems

Figure 7.1 represents a ridge and a groove and the direction of the x and z axes. These systems are deemed infinite in the direction perpendicular to the page. In all calculations throughout this chapter, the width of all defects is

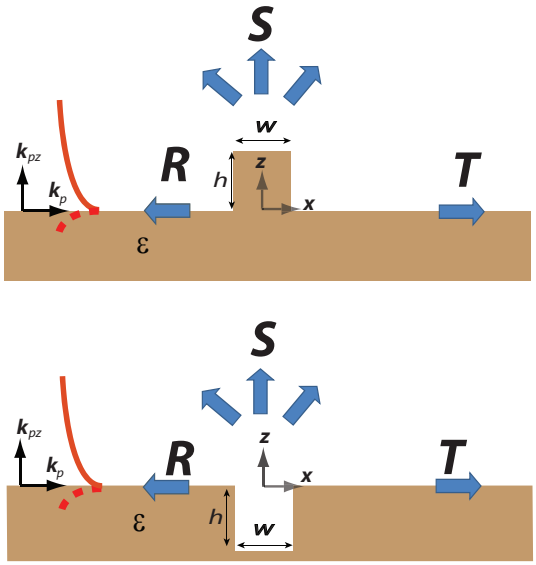

Figure 7.1: Schematic representation of the two scattering systems: ridges(protrusions) and grooves (indentations). A surface plasmon polariton impinges at normal incidence. Part of the incident energy flux of surface plasmon is transmitted, while another part is scattered in the reflection and out-of-plane radiation channels. In this chapter the width (w) of the defects is set to $w = 100$nm.

fixed to the value of $w = 100$nm (which is experimentally viable). First, we shall consider individual defects of several heights: from shallow defects to deep defects. Secondly, we shall consider arrays of ridges and grooves consisting of identical defects periodically distributed and, again, while the individual width of the defects is kept equal to $w = 100$nm, the depth of the defects will be varied. Notice that defects are made on a thick metal slab which we approximate by an air-metal semi-infinite space. The considered material is silver and its dielectric constant is taken from Ref.[61]. Absorption is taken into account because the size of the defects may be comparable with the SPP absorption length. The solutions are attained through the Green tensor approach, which is a standard theoretical method to solve electromagnetic scattering problems [3, 5–7, 10, 12, 13, 93].

Ridges and Grooves will be illuminated by a surface plasmon at normal incidence, propagating in the x-direction with in-plane wavevector[43, 94] $k_p = (2\pi/\lambda)\sqrt{\varepsilon\,(\varepsilon+1)^{-1}}$, where λ is the free-space wavelength.

The scattering problem is defined in terms of the fraction of the SPP energy flux carried by Transmission (T), Reflection (R) and out-of-plane Radiation (S). For economy of language, throughout the chapter we shall refer to the out-of plane radiation of the impinging SPP energy flux simply as radiation. Unless otherwise stated, transmission, radiation and reflection shall be represented as a function of wavelength from a short-wavelength edge of visible light at $\lambda_1 = 500nm$ to a long-wavelength edge of near-infrared light $\lambda_2 = 1000nm$. The SPP has a penetration in the air semi-space determined by $(\Im\{k_{pz}\})^{-1}$, where $k_{pz} = 2\pi/(\lambda\sqrt{\varepsilon+1})$ is the vertical component of its wavevector. The SPP penetration in air grows monotonically from visible to infrared wavelengths, from the value of $(\Im\{k_{pz}\})^{-1} \simeq 200nm$ at λ_1, to the value of $(\Im\{k_{pz}\})^{-1} \simeq 1080nm$ at λ_2. Approximately at $700nm$, in the border between optical and infrared wavelengths: $(\Im\{k_{pz}\})^{-1} \simeq 480nm$

7.3 Individual Defects

7.3.1 Individual Ridges

Figure 7.2 renders the evolution of the SPP transmission spectrum along with the SPP reflection and radiation spectra, of an individual ridge as its height is varied from 25nm to 500nm.
The scattering by shallow defects ($h << \lambda$) has been previously analyzed in Chapter 6. In this case the size of the defect is much smaller that the free-space wavelength of the incident light and, therefore, the defect scattering can be associated to the emission of a point-dipole. In such case we have Rayleigh scattering[58]. In 2D systems, such as shallow ridges ($h = 25nm$), R and S exhibit a smooth Rayleigh-type decay with wavelength, that scales as λ^{-3}. Transmission for shallow defects, undergoes a rapid increase with an increasing wavelength. Nevertheless Fig.7.2(a) shows that, as h varies from $h = \lambda_1/10$ to $h = \lambda_1$, transmission gradually drops.

Correspondingly, as shown in Fig.7.2(b), the increase in h results in a monotone increase in reflection until the ridge is almost a perfect reflector, reaching its maximum efficiency of 90% of the incident plasmon (including absorption) at a height such that: $h > (\Im\{k_{pz}\})^{-1}$, at optical wavelengths, in a neighborhood of λ_1. This behavior is to be expected: when the height of the ridge is greater than the SPP penetration in the air, the ridge acts as a barrier reflecting the SPP almost completely. Notice the ridge reflection is always maximum in a neighborhood of the short-wavelength edge of the spectrum λ_1.
Figure 7.2(c) shows the evolution of the radiation spectrum for the same set of heights. When the ridge is still shallow ($h = \lambda_1/10$), a peak in radiation is formed at λ_1 where $h < (\Im\{k_{pz}\})^{-1}$. This peak is red-shifted at every height increment considered. Finally when the ridge height is $h = 500nm$ and $h \geq (\Im\{k_{pz}\})^{-1}$ in the optical range (up to 700nm), the radiation peak is red-shifted to near-infrared wavelengths $\lambda > \lambda_2$. For all values of h considered, the peaks in the radiation spectrum are smaller than the maxima in the reflection spectrum.
Figure 7.3 represents an extremely tall ridge ($h = 1.5\mu m$). This is considered merely for academic reasons as it contains information on the main scattering channels in different regions of the spectrum. Notice that, this time we are representing the spectra of T, R and S from a wavelength of 500nm to 4500nm. In this case, reflection is the the strongest scattering channel in the optical region. Radiation, however, is the main scattering mechanism in the interval [1500 − 2000nm] of the spectrum. Even then, and unlike the situation in the reflection channel, radiation is not systematically enhanced with the ridge height. In fact, its scattering efficiency never overcomes 50% of the

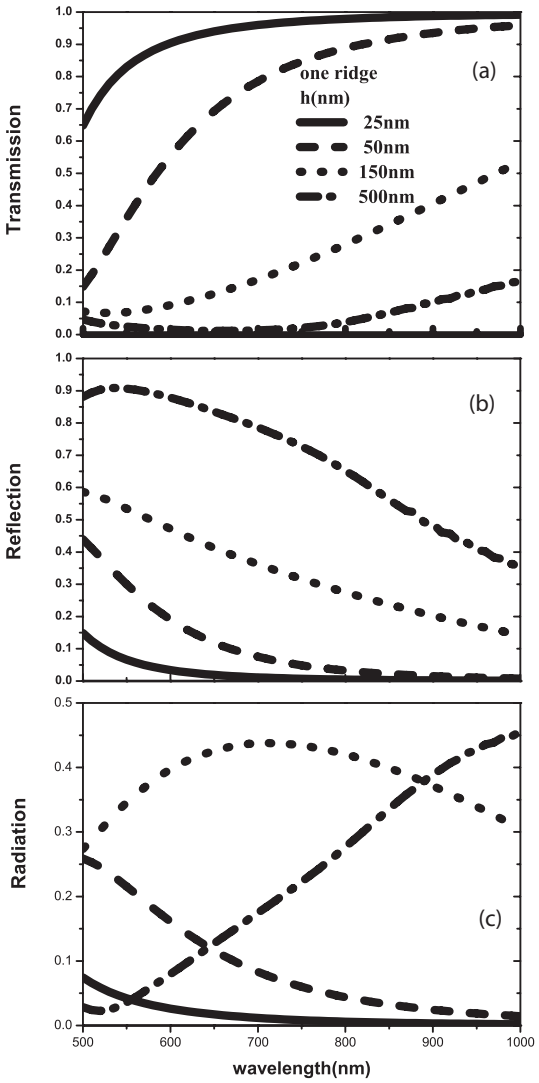

Figure 7.2: The transmission (panel (a)), reflection (panel (b)) and radiation (panel (c)) of a surface plasmon, as a function of wavelength, produced by *one ridge* at different heights. The ridge width is fixed to $w = 100$nm while its height h is varied from 25nm to 500nm. The material is silver.

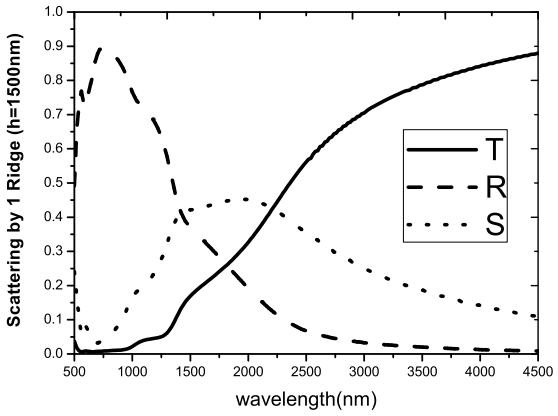

Figure 7.3: The transmission (T), reflection(R), and radiation(S) of a surface plasmon by one ridge as a function of wavelength. The ridge width is $w = 100$nm while its height is $h = 1500$nm. The material is silver.

incident energy flux.

7.3.2 Individual Grooves

The surface plasmon scattering by a groove has been studied much more than the one by a ridge, see for example Ref.[68, 85, 86, 95, 96]. In a simple model[91, 97] the field excited in the groove by the incident SPP can be expressed as a superposition of the waveguide modes. The excited modes propagate in the *vertical direction*. Depending on the depth of the groove its interaction with the plasmon has two extremes. Maximum scattering may occur or, conversely, the groove may become almost invisible to the incident plasmon and the defect can be seen as an impedance defect[93]. In the first case the groove is resonant. In the second case transmission is maximum and the total scattering (R+S) is minimum (practically zero).
We have found that for the depths rendered in Fig.7.4, a resonant behavior is observed in the spectrum of grooves. Conversely within the range of depths, from $h = 150$nm to $h = 200$nm, between 70% and 90% of the incident plasmon is transmitted.
In Fig.7.4 we start from the case of a shallow groove, which we have analyzed in Chapter 6. In this case the groove is extremely shallow and only exhibits a Rayleigh-type decay with wavelength. Figure 7.4 shows that, for $h \simeq \lambda_1/10$, a resonant peak appears close to λ_1. Let us label this resonance by the integer $n = 1$. If we increase the depth to $h = \lambda_1/5$, the groove resonance $n = 1$ is red-shifted as well as damped, since at longer wavelengths transmission tends to increase [97].
At a critical depth of a half-wavelength ($h = \lambda_1/2 = 250$nm) Fig.7.4 exhibits a new resonant wavelength close to λ_1, let us label it $n = 2$. When the groove depth is further increased to $h = 450$nm, the $n = 2$ resonance is red-shifted, at $\lambda \simeq 850$nm, and damped. At the same time the depth of the groove is large enough to meet a new resonant condition ($n = 3$) in the neighborhood of λ_1. For greater depths this behavior culminates in a multiple-resonance pattern, represented by a succession of damped and broadened oscillation between points of maximum scattering and maximum transmission as in Fig.7.5.
In order to show the effect of absorption on the system Fig.7.5 presents the spectra for T, R and S, by a 1μm deep groove with and without absorption. The figure shows that the spectral position of both the resonant wavelengths and the transmission maxima are the same in the loss-free and lossy case. Furthermore, in both cases, the net scattering (reflection+radiation) is zero at wavelengths of transmission maxima. The effect of absorption is to attenuate the amplitude of the reflection and radiation peaks at resonant wavelengths as well as those of the transmission maxima (which in the loss-free case give T=1).
As it turns out, the resonant coupling of the groove and the plasmon results in both radiation and reflection occurring at the same wavelength. Yet, we find

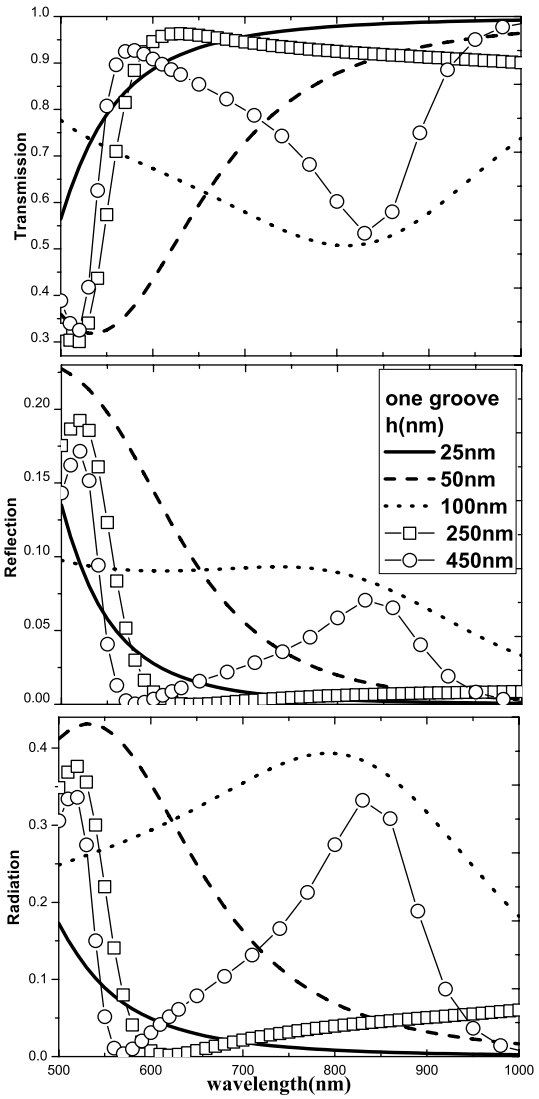

Figure 7.4: The transmission (panel (a)), reflection (panel (b)) and radiation(panel (c)) of a surface plasmon, as a function of wavelength, produced by *one groove* at different depths. The groove width is fixed to $w = 100$nm while its depth h is varied from 25nm to 450nm. The material is silver.

that radiation is the most efficient scattering mechanism in a groove especially at long wavelengths, as seen for example in Fig.7.4 and Fig.7.5.
The assumption that the field propagates in the z-direction implies that the relation between wavelength and height at the critical points of maximum and minimum scattering (or maximum transmission) is: $h = a\lambda + b$. Neglecting the effect of the groove width on the resonant conditions and considering that the groove is defined by walls made of perfect electrical conductor (PEC), the values for a are[91, 97]: $a = a_{max} = (2n+1)/4$ and $a = a_{min} = n/2$ for for maximum and minimum scattering respectively, while $b = 0$ in both cases.

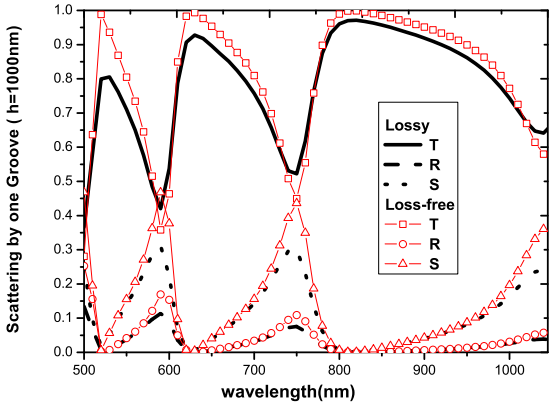

Figure 7.5: The transmission (T), reflection(R), and radiation(S) of a surface plasmon by one groove as a function of wavelength. The groove width is $w = 100$nm while its height is $h = 1000$nm. The material is silver.

Our exact calculations are in accordance with the linear dependence between h and λ at maximum and minimum scattering, as shown in Fig.7.6. We have made linear fits on all resonance and transmission curves in Fig.7.6, calculated the relative error on them being straight lines (defined as the ratio of the standard error on the slope with the fitted slope). The relative error is always smaller than 4% in the optical range. In the Table below we have tabulated the slopes and intercepts of the linear fits of our results, for the scattering maxima and minima.

n	a_{min}	b_{min}	a_{max}	b_{max}
1	0.49	-60.8	0.188	-51.
2	0.89	-68.7	0.63	-74.
3	1.37	-128	1.05	-91.4
4	2.	-260	1.48	-113.

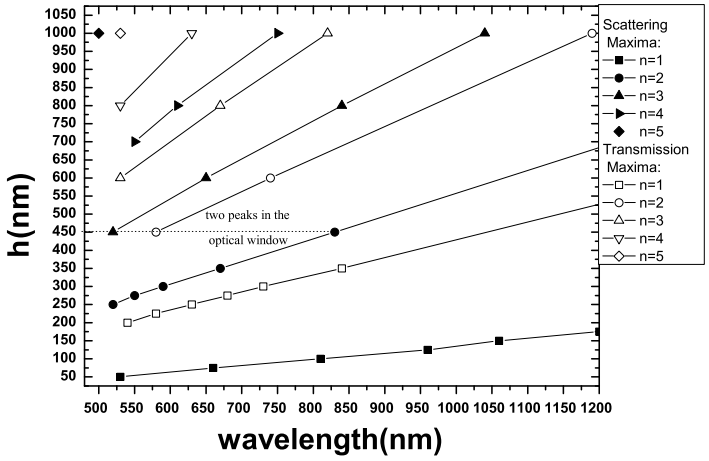

Figure 7.6: The depths (h), of one groove of width $w = 100$nm, against wavelengths providing either maximum scattering or maximum transmission of SPPs. The maxima, of both transmission and scattering, that appear on the same straight line are labeled by an integer index n.

The exact results are consistent with the model of Ref.[91] but show deviations from it. This can be associated to the penetration of the field in a real metal which implies that the field inside a groove has a small wavevector component parallel to the surface $k_x \neq 0$. Thus while in a PEC $k_z = 2\pi/\lambda$, in a real metal k_z is larger. Correspondingly, the condition of resonant depth occurs at smaller values than those predicted by the PEC approximation. This is reflected in the smaller values obtained for the coefficients a_{min} and a_{max}.

7.4 Individual Ridges vs. Individual Grooves: reflection and radiation

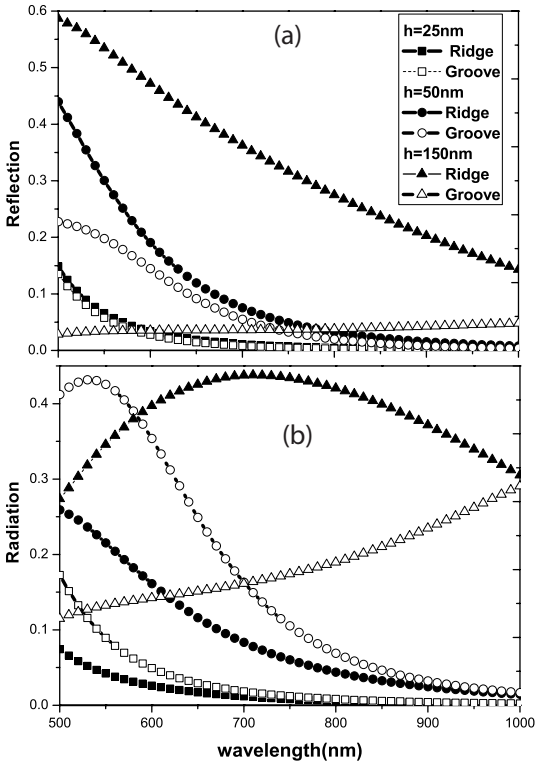

Figure 7.7: The reflection (panel(a)) and radiation (panel(b)) of a surface plasmon polariton by both one ridge and one groove as a function of wavelength, for different heights/depths (h). The defects width is $w = 100$nm while their heights are $h =$ 25nm,50nm and 150nm. The material is silver.

So far we have seen that the ridge capability to reflect and radiate SPPs increases systematically with height. By contrast in grooves, reflection and radiation of SPPs occurs mainly at resonant depths.
In the Rayleigh limit the comparison between a ridge and a groove, revealed two limiting cases (see Chapter 6). *i*) When the defects are very small and square, the reflection between ridges and grooves of the same size is comparable, while the out of plane radiation is always larger in a groove. *ii*) When defects are shallow but very long in the x-direction ridges and grooves have similar scattering. Yet, this requires needle-type defects, with $w >> h$ (at least $w > 10h$). As the aspect ratio (w/h) of the defect is varied we pass gradually from the case *i*) to the case *ii*).
The scattering coefficients of the shallow defects considered ($w = 4h$), can be qualitatively associated with case *i*) as shown in Fig.7.7. The rest of the

section is devoted to studying how this scenario changes when the defects are not shallow.

7.4.1 Reflection of surface plasmons

Figure 7.7(a) compares the reflection spectra of ridges and grooves of the same sizes, as h is increased up to $h \sim \lambda_1/3$. The figure illustrates the greater reflection efficiency of ridges over grooves, for this range of values of h. Ridges can efficiently reflect the incident SPP, even when their height is only a few tens of nanometers. Figure 7.7(a) shows that a 50nm-tall ridge ($h \sim \lambda_1/10$) is able to reflect almost 45% of the incident SPP. We find that, in agreement with the results of Ref.[95], a groove never overcomes a reflection efficiency of 30% of the impinging SPP energy flux. For larger values of h than those rendered in Fig.7.7(a), the ridge maximum reflection efficiency is much greater than that of a groove with the same size and shape.

7.4.2 Out-of-plane Radiation of surface plasmons

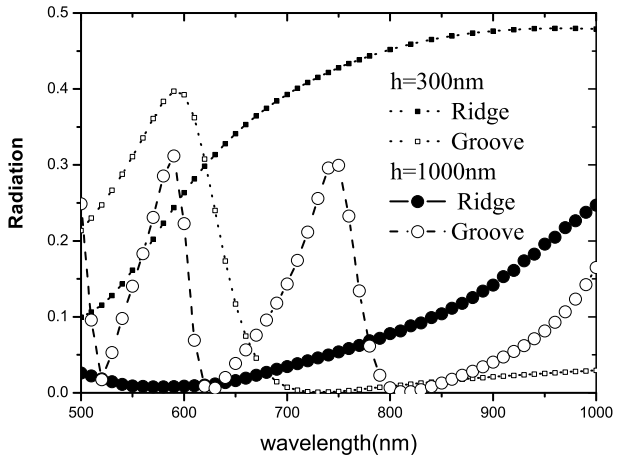

Figure 7.8: The radiation of surface plasmon polaritons by both one ridge and one groove as a function of wavelength, for different heights/depths (h). The defects width is $w = 100$nm while their heights are $h = 300$nm and $h = 1000$nm. The material is silver.

The comparison between the radiation efficiency of SPPs by grooves with the one by ridges of the same size, is more complex.
Figure 7.7(b) shows that, the groove is resonant at $h = \lambda_1/10 = 50nm$. In a neighborhood of the resonant region, at optical wavelengths, the groove gives

greater overall radiation than a ridge with the same size. However when the groove is non resonant ($h = 150$nm), throughout the spectral region $[\lambda_1, \lambda_2]$, its radiation of SPPs is smaller than the one by the ridge, in the entire spectral region.
The second resonance of the groove emerges at $h = 300nm \sim \lambda_1/2$. As seen in Fig.7.8 such groove resonance presents a localized peak over a relatively small spectral region in the visible range. Instead for a ridge of the same height, the SPP radiation grows monotonously in the range $[\lambda_1, \lambda_2]$. Figure 7.8 also shows what happens eventually for very deep defects. The maximum of the ridge radiation spectrum red-shifts and occurs at $\lambda > \lambda_1$. The radiation spectrum of a groove tends to become prominent in the optical region, presenting a multi-resonant pattern of emission lines.

In conclusion, an individual ridge scatters the energy flux of an incident SPP, more efficiently than a groove of the same height. As the ridge gets tall radiation becomes the main scattering channel at optical wavelengths, while radiation is red-shifted. Its radiation maxima occur at different wavelengths from reflection maxima. However a single groove presents resonances that depend on the groove depth. Its resonant radiation and reflection peaks occur at the same wavelength.

7.5 Arrays of surface defects

We shall now highlight the effects exhibited collectively by ridges and grooves, which do not appear in individual defects, and then comment briefly on the cases in which scattering is explainable in terms of the individual behavior of defects.
As known, a set of scatterers periodically distributed, exhibits photonic band-gaps[98, 99]. We have considered arrays of five *identical defects* of width $w = 100nm$, with a periodicity of 600nm, *and with height h, which is varied.* As it turns out, this is a sufficient number of defects to observe band-gaps effects. The periodicity is chosen to produce band-gap effects within the optical range. Most of the phenomenological analysis, for our scattering system, is based

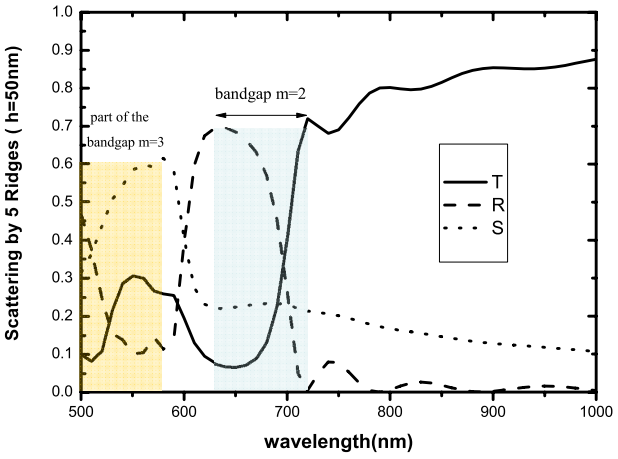

Figure 7.9: The transmission (T), reflection(R), and radiation(S) of a surface plasmon by an array of five ridges, and period 600nm, as a function of wavelength. Each ridge has width $w = 100$nm and height $h = 50$nm. The material is silver. The spectral position of the band-gaps is indicated.

on the concept of band-gap structure. The position of the band-gap short-wavelength edge $\lambda_-^{(m)}$ is determined by the Bragg law[69, 99] $k_p d = m\pi$. The width of the band-gap spans from $\lambda_-^{(m)}$ to $\lambda_+^{(m)}$. Former results[68, 92] have shown that reflection has a peak at $\lambda_-^{(m)}$ while radiation has a peak at $\lambda_+^{(m)}$. The short-wavelength edge of the band-gap $m = 2$ is at about $\lambda_-^{(2)} \approx 630nm$. This is the main band-gap observable in our system within the spectral range $[\lambda_1, \lambda_2]$. Yet the $m = 3$ bands-gap will be part of the discussion.

7.5.1 Ridges arrays: Collective effects

For the chosen period, ridge arrays produce band-gaps that span a region within $[\lambda_1, \lambda_2]$. In Fig.7.9 we can notice the whole $m = 2$ band-gap, $[\lambda_-^{(2)}, \lambda_+^{(2)}]$,

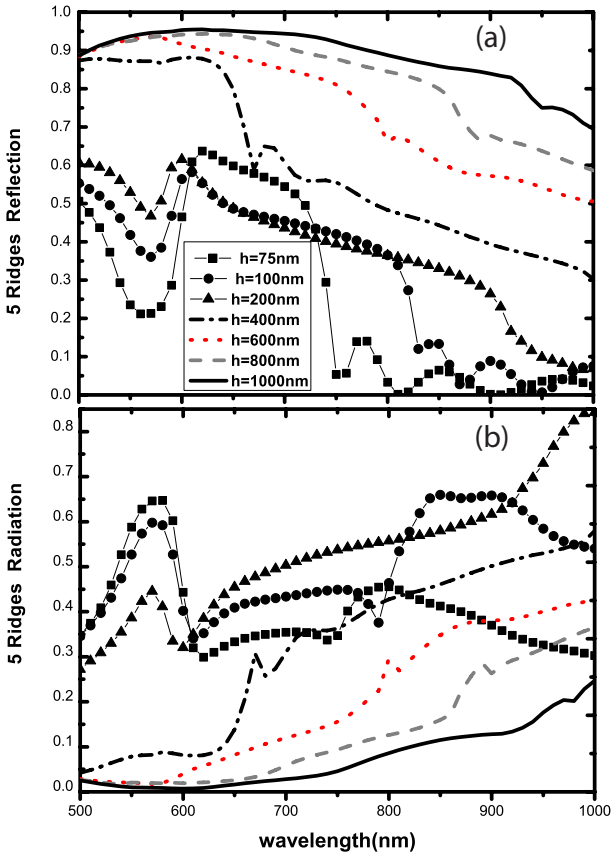

Figure 7.10: The reflection (Panel(a)) and radiation (Panel(b)) of a surface plasmon polariton by an array of five ridges (period 600nm) as a function of wavelength, for different heights (h). Each ridge has width $w = 100$nm while its height is varied from $h = 75$nm to $h = 1000$nm. The material is silver.

as well as the long-wavelength edge $\lambda_+^{(3)}$ of the $m = 3$ band-gap at 550nm. At wavelengths in the neighborhood of $\lambda_+^{(3)}$ the impinging SPP energy is scattered mainly into the radiation channel, while in a neighborhood of $\lambda_-^{(2)}$, it goes mainly into the reflection channel. These two band-gaps are very close together. When the incident SPP free-space wavelength is just outside the $m = 3$ band-gap, $\lambda > \lambda_+^{(3)}$, the main scattering mechanism becomes the build-up to the reflection maximum, up to the wavelength $\lambda_-^{(2)}$. Notice that the $m = 3$ band-gap is opened by fulfilment of Bragg's condition at about 430nm. Yet we have not considered this region because the SPP propagation in silver,

at these wavelengths, is curtailed by large absorption effects.
An increase in the height (h) of the ridges in the array results in band-gap broadening. As the band-gap m=2 gets larger, the reflection peak at $\lambda_-^{(2)}$ becomes more asymmetric. In fact, for a fixed h, the asymmetry in the reflection peak $\lambda_-^{(2)}$ is caused by the progressive growth of the radiation spectrum throughout the gap. The band-gap, and hence the growth of the radiation spectrum, spans from $\lambda_-^{(2)}$ to $\lambda_+^{(2)}$. If, in a different configuration, $\lambda_+^{(2)}$ is shifted to a longer wavelength $\lambda_+'^{(2)}$, the related radiation growth is also extended to $\lambda_+'^{(2)} > \lambda_+^{(2)}$, resulting in a more asymmetric reflection peak at $\lambda_-^{(2)}$. *Therefore the asymmetry of the reflection peak is also associated to the spectral width of the gap.* Figure 7.10(a) shows that if h is increased from $h = 75$nm to $h = 200$nm, the resulting band-gap expands its width to include a broader spectral region.
Notice that the reflection spectra in Fig.7.10(a), exhibits a set of small multiple-resonance peaks at $\lambda > 750$nm, for $h = 75$nm and $h = 100$nm . Correspondingly, radiation also exhibits such small peaks, in the near-infrared region of the spectra rendered in Fig.7.10(b). The small peaks are array finite-size effects formerly discussed in Ref.[90]. Notice that, as explained in Ref.[90], finite-size effects occur outside the band-gap edges. Figure 7.10(b) also shows that, as h is raised, the band-gap spectral width is broadened. As h varies from 75nm to 200nm, the position of $\lambda_+^{(2)}$ is shifted to from 750nm to 1000nm.
While band-gap effects are shifted to infrared wavelengths beyond $1\mu m$, our system at $h = 200nm$ is at a critical height in which its band-gap structure begins to become weaker, at optical wavelengths, in the neighborhood of $\lambda_+^{(3)}$. In fact as h approaches 200nm, it becomes comparable to $\Im\{k_{pz}\}^{-1}$ in the spectral interval $[500-600]$nm. As a result the radiation peak, that had appeared in Fig.7.9 at $\lambda_+^{(3)}$ drops in Fig.7.10. As seen in the figure reflection increases in the spectral interval $[500-600]$nm, at the same rate as the radiation peak at $\lambda_+^{(3)}$ tends to vanish.
For $h = 400$nm the height of the ridges is such that $h \sim \Im\{k_{pz}\}^{-1}$, in the range $\lambda \in [500-600]$nm. At these wavelengths such arrays consist of ridges so tall that the first one or two block most of the impinging plasmon. As seen in Fig.7.10 when h is increased from 400nm to 1000nm, reflection rapidly becomes the only scattering channel in the range $[\lambda_1, \lambda_2]$ while radiation vanishes. Therefore, as h approaches 1000nm, band-gap effects tend to disappear in the whole optical spectrum and the scattering of the incident SPP by the ridge array, can be interpreted, mainly, as the individual scattering of the first one or two ridges.

7.5.2 Grooves arrays: Collective effects

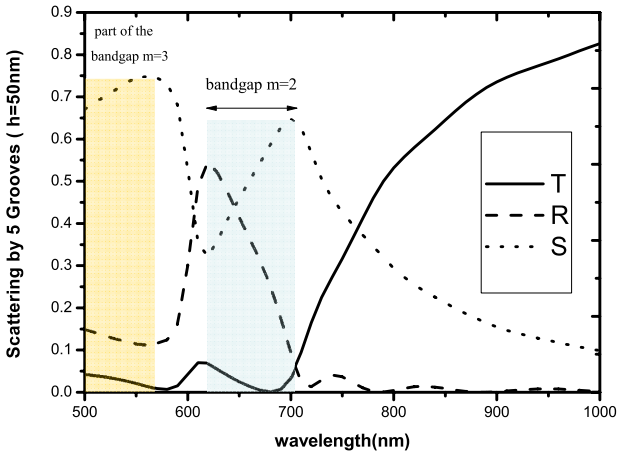

Figure 7.11: The transmission (T), reflection(R), and radiation(S) of a surface plasmon by an array of five grooves, and period 600nm as a function of wavelength. Each groove has a width $w = 100$nm and depth $h = 50$nm. The material is silver. The spectral position of the band-gaps is indicated.

As found for individual grooves, the scattering coefficients of grooves arrays have an oscillatory behavior that depends on the depth of the grooves and the free-space wavelength of the incident SPP. Several works [91, 97, 100] have studied SPP scattering by groove arrays. In order to compare with the ridge array analyzed before, in the following, we shall consider the SPP scattering by 5 grooves with period 600nm and width 100nm, and different depths.
First of all in Fig.7.11 we present the exact result for the array of grooves considered in Ref.[68] with an approximate method. The two results are extremely similar, except that the radiation growth within the band-gap is much more evident in the exact result than in the approximate one. Noticeably Fig.7.11 features the same radiation peak at $\lambda_+^{(3)}$ as Fig.7.9.
As opposed to the individual behavior of a resonant groove, in groove arrays the reflection and radiation maxima are effectively decoupled, when the grooves are interacting collectively. In fact, in this case the reflection peak is imposed at $\lambda_-^{(m)}$ by Bragg's interference, based on the period of the array. However, since the radiation peak is caused by the decrease of the reflection (and transmission) within the gap the two peaks appear at different wavelengths, unless the spectral width of the band-gap is zero.
When, in the same configuration, the depth of the grooves is increased up to $h = 200$nm, the interaction between the incident SPP and the groove weakens.

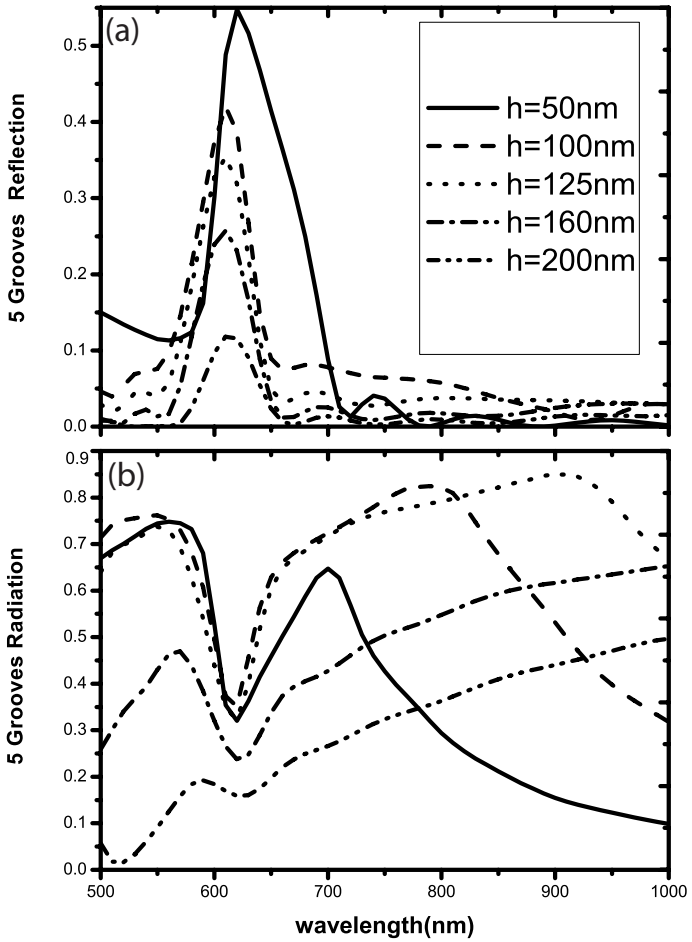

Figure 7.12: The reflection (Panel(a)) and radiation (Panel(b)) of a surface plasmon polariton by an array of five grooves, and period 600nm, as a function of wavelength, for different depths (h). Each groove has width $w = 100$nm while its depth is varied from $h = 75$nm to $h = 1000$nm. The material is silver.

As a result the related reflection maximum is systematically reduced, as seen in Fig.7.12(a). Figure 7.12(b) shows that the radiation maximum at $\lambda_+^{(3)}$ disappears. Similarly, as h increases, the radiation also vanishes throughout the spectrum $[\lambda_1, \lambda_2]$. The radiation peak at $\lambda_+^{(2)}$ is only observable for $h = 50$nm and $h = 100$nm, while for deeper grooves it is red-shifted beyond $\lambda > \lambda_2$.

Figure 7.13 represents the result for very deep grooves. Note the small re-

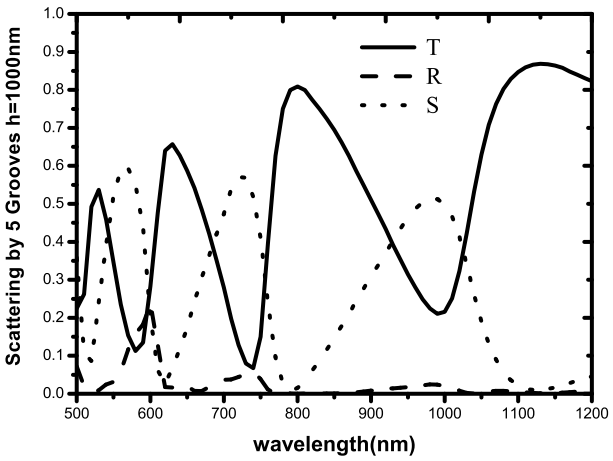

Figure 7.13: The transmission (T), reflection(R), and radiation(S) of a surface plasmon by an array of five grooves, and period 600nm, as a function of wavelength. Each groove has a width $w = 100$nm and depth $h = 1000$nm. The material is silver.

flection peak at $\lambda_-^{(2)}$ is the only one decoupled from radiative emission lines. Comparison with Fig.7.5 suggests that the interaction between the 5 grooves in Fig.7.13 is weak and that the scattering produced by the array is rather a superposition of the individual behavior of each groove. The comparison between the two figures also shows that the values of radiation peaks are about 60% in Fig.7.13 and about 30% in Fig.7.5. That is, at resonance, the fraction of impinging SPP flux radiated by five grooves doubles the one radiated by a single groove, of the same size. Besides radiating more efficiently, five grooves produce more dissipation of the impinging energy flux than one does, and this causes less overall transmission throughout the spectrum in Fig.7.13.

Finally notice that, as found for individual grooves, the out-of-plane radiative loss is the dominant scattering mechanism for groove arrays. However arrays of grooves can produce considerable reflection of surface plasmons.

7.6 Arrays of Ridges vs Arrays of Grooves

As seen, ridge arrays produce larger reflection efficiencies than groove arrays, even when the latter are resonant.
Comparing the increase of the radiation peak at $\lambda_+^{(2)}$ in Fig.7.11 with that in Fig.7.9 suggests that groove arrays are better radiative emitters when the defects are shallow. Accordingly, in groove arrays the main scattering mechanism within the band-gap is radiation. In ridge arrays the main scattering mechanism within the band-gap is reflection. The spectral width of the band-gap produced by either ridges or grooves of $h = 50$nm is practically equal. Note that grooves are not resonant within the band-gap $[\lambda_-^{(2)}, \lambda_+^{(2)}]$ for $h = 50$nm (see Fig.7.6). However, for $h = 100$nm grooves are resonant within the spectral range $[\lambda_-^{(2)}, \lambda_+^{(2)}]$. In this case the band-gap width produced by the groove array is larger than the one produced by the ridge array. In fact, as seen in Fig.7.12(b), $\lambda_+^{(2)}$ in grooves is close to 900nm, while as seen from Fig.7.10, $\lambda_+^{(2)}$ in ridges is close to 800nm.
Deep groove arrays undergo larger dissipative loss than tall ridges, and their maximum out-of-plane scattering efficiency is about 60%. Tall ridges dissipate less of the SPP energy, achieving a maximum reflection efficiency over 90%.

7.7 Conclusions

We have studied the individual and collective scattering of ridges and grooves in the optical range. The width of the defects was always fixed to a typical value of 100nm and the period in arrays was fixed to 600nm, while the height and depths were varied from 25nm to about one micron. To the best of our knowledge this is the first comparative treatment between ridges and grooves where their depth is systematically increased. We found ridges are very good reflectors, featuring (90%) reflection efficiency, including absorption. The related reflection becomes the main scattering channel in the optical range as the ridge height is increased, while radiation is red-shifted to infrared wavelengths. The reflection and out-of-plane radiation maxima are found at different wavelengths. Ridges produce more scattering than groves in general, but the latter are more versatile. In fact, by adjusting their depth to the free space wavelength, we can produce a tunable resonance or virtual invisibility. At resonance grooves exhibit radiation and reflection at the same wavelength, but these can be decoupled through band-gap effects. Reflection is the least efficient mechanism for both individual and collections of grooves. In both cases the reflection peaks tend to disappear in the long-wavelength limit.
In shallow arrays (h=50nm) ridges and grooves have similar band-gap structures. However within the gap ridges can reflect incident SPPs, more efficiently than grooves and, in turn, grooves can radiate the incident SPP more efficiently than ridges. An array consisting of grooves whose depth is resonant at wavelengths within the band-gap, produces a larger band-gap spectral width than an array of ridges with the same size and period, in the considered configurations.
A SPP suffers large dissipative losses when scattered by a deep groove array, and the maximum fraction of the impinging SPP energy flux scattered by a groove array is 60%. A tall ridge array can reflect more than 90% of the impinging SPP energy flux.

Chapter 8

Scattering of surface plasmon polaritons by 3D surface defects

Surface plasmon polaritons (SPPs), i. e., surface electromagnetic excitations propagating along metal-dielectric interfaces,[43] are currently attracting a great deal of attention.[41, 79] One of the most interesting aspects of SPPs is the possibility of concentrating and guiding electromagnetic radiation at a sub-wavelength scale by using surface nanostructures. This possibility has been intensively explored in recent experimental and theoretical investigations for a range of nanostructures. SPP wave guiding properties of metallic stripes having widths within the micrometer range have been studied both theoretically [101, 102] and experimentally,[103–105] involving SPPs propagation along stripes of finite length and arrangements of narrow slits and indentations.[68, 70, 81, 106–110] Another type of SPP-based micro-optical components can be realized by making use of individual metallic nanoparticles arranged to form various structures such as linear chains or two-dimensional arrays. Experimental studies[76, 111] showed that nanoparticle ensembles on metal surfaces can be used to create efficient micro-optical components for SPPs, such as mirrors, beam splitters and interferometers. SPP scattering plays the main role in electromagnetic interactions occurring in ensembles of surface nanoparticles used for realization of SPP micro-components. In general, the corresponding scattering problems are very complicated. One of the possible approaches is to make use of the point-dipole approximation (PDA)[30, 112, 113] that has successfully been applied to model a number of SPP micro-components.[114, 115] In general, the PDA might fail to properly represent even a very small particle,[116, 117] so that one should solve the scattering problem exactly or use other simplifying assumptions. However the theoretical investigation of SPP scattering processes on various objects is a

very intricate problem that requires elaborate numerical calculations even in the relatively simple case of an individual symmetrical scatterer. In this chapter we present general theoretical considerations of the SPP scattering and attain expressions for the differential and total scattering cross sections for the SPP scattering by a finite-sized nanoparticle placed near a planar metal surface. The scattering process includes elastic scattering of the incident SPP into SPPs propagating in different directions and its (inelastic) scattering into field components propagating away from the surface, as well as radiation absorption by the (metal) nanoparticle. Special attention is paid to the connection between the general treatment and the PDA. We calculate the SPP extinction cross sections for a gold nano-cube placed near a gold planar surface and analyze the dependence of the extinction spectrum on the size of the cube and on its position with respect to the gold surface. In this case, the dielectric function of the metal in the planar metal-dielectric interface system is assumed to have a negligible imaginary part, while the permittivity of the scatterer has an arbitrary value. Since the PDA is widely exploited for modeling of SPP scattering, we compare the SPP extinction spectra calculated for a finite-sized cubic scatterer with those obtained using the PDA. This comparison clarifies a general problem encountered when applying the PDA, which was found to provide only qualitative agreement with experimental results. Throughout this chapter, we use the gold permittivity data from Ref. [61] unless otherwise stated.

8.1 Numerical consideration and comparison with Point Dipole Approximation

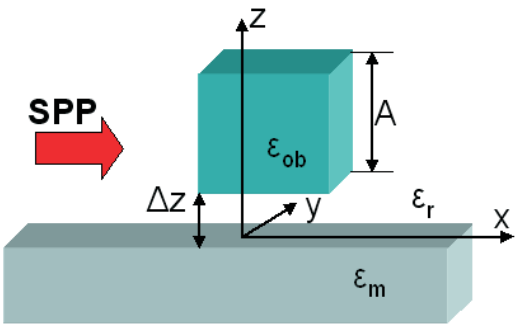

Figure 8.1: (Color online) Schematic representation of a scattering system: an SPP monochromatic wave is scattered by a particle with dimension A and dielectric constant ε_p. Δz is the particle-to-surface gap

Let us consider SPP scattering by a single gold particle located near a gold surface. Recently, the electric EM field distribution of a related system (a non-cubic particle placed on a thin metal film, illuminated by a plane wave) has been analysed[118, 119]. Here we consider on the SPP scattering by a cubic particle with side A, first with the exact GTA and then we compare the results with Point Dipole Approximation (PDA). In the latter we approximate the polarizability of the cube on the surface with that of a sphere in vacuum. In our study, the particle is separated from the surface by a finite space gap Δz (Fig.8.1), defined such that if the particle touches the surface $\Delta z = 0$.
In the numerical procedure (with no approximation) the value of discretization step is determined by the convergence of the solution and depends on the wavelength. When the wavelength is non-resonant (see below) the discretization step has been taken equal to $A/10$, otherwise it was necessary to take a mesh size equal to $A/20$. The criterion was that a finer discretization step would not practically produce a difference from the value of the cross-section obtained in the previous cases (with a precision of two significant figures).
Let us first consider the case $\Delta z = 0$. Figure 8.2 represents the spectrum of SPP extinction efficiency (σ_{ext}/A) for different sizes of the scattering cube. The resonance peak corresponding to the localized surface plasmon (LSP) mode of the cube appears in the considered wavelength range for relatively small cubes. This mode is excited by the electric field of the incident SPP wave. Detailed discussion of the plasmon eigenmodes for a cubiclike particle in free space can be found in Refs. [120, 121]. Note that the field inside the small cubes (in the quasistatic limit) is determined by the induced electronic charge that is not distributed homogeneously on the surface. As a result the multipolar

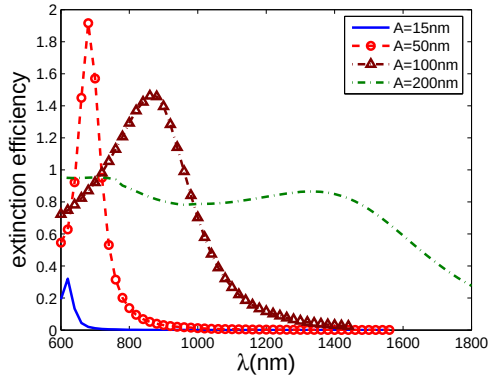

Figure 8.2: (Color online) Spectrum of SPP extinction efficiency σ_{ext}/A, calculated using the GTA, for gold cubic particles with side dimension A. The particle-to-surface gap is $\Delta z = 0$.

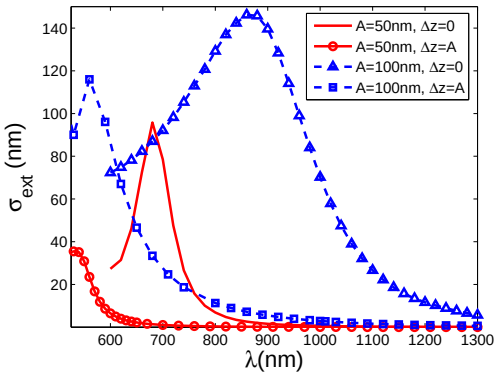

Figure 8.3: (Color online) Spectrum of SPP extinction cross section σ_{ext}, calculated using the GTA, for gold cubic particles with side dimension A for different particle-to-surface gaps Δz.

charge distributions always exist independently of the cube size. However the high multipolar LSPs correspond to smaller wavelengths than dipolar one[122]. When the quasi-static field gives the main contribution in the total field inside the particles (curves for A=15 nm and 50 nm), the resonance extinction efficiency increases as the size of the particle increases, whereas for relatively large particles this value decreases as the size of the particle increases (Fig.8.2). At the same time for the large particles (A=100nm and A=200nm) the peak is broadened because of the contributions of SPPs and of radiation from the total electric field inside the cube. Furthermore, increasing the size of the particle will redshift the resonance wavelength. This behaviour is similar to the case of light being scattered by a metal nano-particle in homogeneous space[123]. However some features related to the presence of the metal surface

can be highlighted. For the small particles ($A = 15$ nm and $A = 50$ nm) the spectrum difference between the resonance position is equal to approximately 100 nm, whereas for larger particles ($A = 100$ nm and $A = 200$ nm) this difference is increased to 500 nm (see Fig.8.2). As a consequence, one may suppose that the red-shift, and hence the interaction of a SPP scatterer with a metal surface, is larger for large particles than for small ones. This is confirmed by the results presented in Fig.8.3. Comparing the extinction spectra for cubes of the same size located on different distances from the metal surface two features appear in Fig.8.3: first, the red shift of the LSP resonance due to the interaction between the scatterer and the metal surface increases with the size of the cube and, secondly, the resonance value of extinction cross section is enhanced, especially for the small particle when the particle approaches the surface and touches it ($\Delta z = 0$). Our calculations have also shown that no significant resonance shift occurs due to the interaction with the surface with SPP, when Δz is larger than $A/2$.

Let us now compare the SPP extinction spectra calculated with the GTA for a cubic gold particle placed in the vicinity of a flat gold surface with the results obtained using the PDA when the scatterer is approximated by a spherical particle of the same volume. In the PDA, the spherical particle corresponding to the cube with side dimension A is described by the polarizability tensor:

$$\hat{\alpha}_d = \frac{3\varepsilon_b}{\varepsilon_p + 2\varepsilon_b}, \tag{8.1}$$

so that the particle is approximate with a point dipole $\mathbf{p} = V_p\,(\varepsilon_p - \varepsilon_b)\,\hat{\alpha}_d\,\mathbf{e}_{spp}(0)$, with V_p is the volume of the cubic particle which determines, the radius of the sphere as $R_p = (3/4\pi)^{1/3}A$ and $z_p = \Delta z + R_p$. For example, for $A = 100$ nm and $\Delta z = A$ one obtains $R_p \approx 62.04$ nm, $z_p \approx 162.04$ nm. The resulting extinction cross sections for 15 and 50 nm cubes are plotted in Fig.8.4a,b for $\Delta z = 0$ and $\Delta z = A$. Respectively, one can see that no resonances appear in the dipole approximation for the considered wavelength range. Moreover, considerable differences exist in all of the represented ranges between the values of the extinction cross sections calculated for finite-sized particles, and within the framework of the point dipole approach, for the particles located just on the metal surface (Fig.8.4a). Note, that the PDA gives a much smaller extinction cross sections than the exact case, not only at resonance but throughout the spectrum. Increasing the distance between the particle and the metal surface hosting SPPs, causes the deviation of the PDA results from the finite size calculations to suddenly decrease (Fig.8.4b) and become negligible under the condition of a large particle-to-surface distance and non-resonant scattering (A=50 nm (Fig.8.4b). It is important to stress that the extinction spectra of the cubic particles located close to the surface are practically parallel to the PDA extinctions for the corresponding spherical particles in those regions

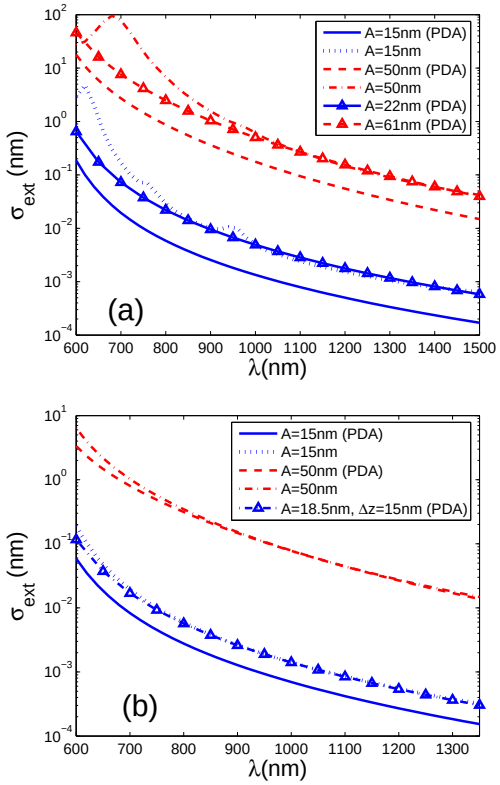

Figure 8.4: SPP extinction spectrum for gold cubic particles with side deimension A and the particle-to-surface gap Δz calculated using both the Green tensor approach (GTA) and the point-dipole approximation (PDA). (a) $\Delta z = 0$: $A_f = 22$nm and $A_f = 61$nm are fitting sizes, which determine the fitting radius $R_p = [(3/4\pi)^{1/3} A_F]$ in the PDA calculations, for cubes A=15nm and A=50nm, respectively.(b) $\Delta z = A$; $A_F = 18.5$nm is the fitting size Radius R_p in the PDA calculation for the cube $A = 15$nm and $\Delta z = 15$nm.

of the spectrum which are far from resonance (Fig.8.4a), and the curves for A=15nm in Fig.8.4b. Only one fitting parameter (the volume or size of scatterer) is sufficient to attain a good agreement between the strict calculation and PDA results for any non-resonant wavelength range (Fig.8.4).

cross section in the point dipole approximation is proportional to V^2, therefore its value increases rapidly on increasing the size of the scatterer. As a result for relatively large particles we have a more complex correlation between the PDA and finite size calculations (Fig. 7). For large scatterers the PDA extinction cross-sections can be larger than the extinction calculated for finite-sized scatterers in the region of relatively small wavelengths. As the size of a scatterer increases the wavelength at which the PDA extinction is equal to the extinction of the finite-sized scatterer is shifted towards longer wavelengths

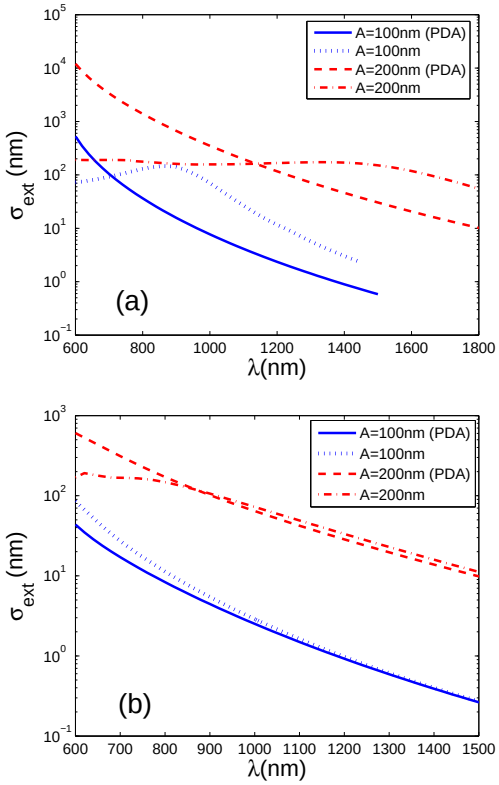

Figure 8.5: SPP extinction spectrum for a gold cubic particle with side dimension A and the particle-to-surface gap Δz calculated using both the Green tensor approach and the point-dipole approximation (PDA). (a) $\Delta z = 0$ (b)$\Delta z = A$.

(Fig. 7a). When the influence of the metal surface on the extinction is small (Fig. 7b) the two approaches give similar results at large wavelengths, where the dipole moment dominates over the rest of the multipole contributions to the total scattering.
Thus the PDA is quite suited for the description of SPP scattering by nanoparticles (cubes) in the case of *weak interaction* between scatterer and metal surface with SPP and far away from the resonance conditions.

8.2 Conclusions

We have analyzed the SPP scattering by a finite-sized nanoparticle placed in the vicinity of a dielectric-metal interface. The SPP extinction spectra for cubic gold particles of different dimensions placed near a flat gold surface have been calculated so as to study the role of finite-size effects on the SPP scattering. It was found that the extinction spectra feature a resonance corresponding to the localized surface plasmon mode mode excited in the cube by the electric field of the incident SPP. The strength and spectral position of the localized surface plasmon resonance strongly depend upon the size of the particle and the particle-to-surface gap. For relatively small particles placed either just on the metal surface or at a small distance from it, comparison between the SPP extinction cross section for cubic particles and that obtained with the PDA showed that the PDA results in significantly smaller scattering cross sections in the considered wavelength range. If the particle-to-surface distance is sufficiently large, the exact calculations and the PDA results were found rather similar in the wavelength range far away from the localized surface plasmon resonance.
Furthermore it was found that, for non-resonant wavelengths and particles being located close to the surface, one can still apply the PDA and obtain good agreement with the exact calculations if the particle volume is used as a fitting parameter in the PDA. Even for relatively large particles, the PDA was found quite accurate once the particle-to-surface distance was sufficiently large. The comparison performed on the two approaches indicates that SPP modeling based on the PDA should be expected to provide only a qualitative agreement with experimental data. However this can be improved by using the particle volume as a fitting parameter. Quantitative agreement requires the usage of accurate calculations, for example based on the GTA, especially for the wavelengths close to the localized surface plasmon resonance. We believe that the presented results should be useful for further understanding and modeling of various SPP scattering phenomena with non-spherical finite-sized particles being used, for example, to realize SPP micro-components.

Appendix A

Appendix

A.1 Unit Vector Identities in Spherical Polar Coordinates

$$\begin{aligned} \mathbf{u}_R &= \cos\phi\sin\theta\mathbf{u}_x + \sin\phi\sin\theta\mathbf{u}_y + \cos\theta\mathbf{u}_z && \text{(A.1)} \\ \mathbf{u}_\theta &= \cos\phi\cos\theta\mathbf{u}_x + \sin\phi\cos\theta\mathbf{u}_y - \sin\theta\mathbf{u}_z && \text{(A.2)} \\ \mathbf{u}_\phi &= -\sin\phi\mathbf{u}_x + \cos\phi\mathbf{u}_z && \text{(A.3)} \end{aligned}$$

$$\begin{aligned} &\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta\, \mathbf{u}_R\mathbf{u}_R = \\ = &\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \left(\sin^2\theta\cos^2\phi\mathbf{u}_x\mathbf{u}_x + \sin^2\theta\sin^2\phi\mathbf{u}_y\mathbf{u}_y + \cos^2\theta\mathbf{u}_z\right) \end{aligned} \tag{A.4}$$

Where we have used the property that any combination of odd powers of sine and cosine gives are null integrals when integrated over

$$\int_0^{2\pi} d\phi \sin\phi\cos\phi = \int_0^{2\pi} d\phi \sin\phi = \int_0^{2\pi} d\phi \cos\phi = 0 \tag{A.5}$$

and consequently off-diagonal components vanish when integrated over the angles. Likewise:

$$\begin{aligned} &\int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta\, \mathbf{u}_\theta\mathbf{u}_\theta = \\ = &\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \left(\cos^2\theta\cos^2\phi\mathbf{u}_x\mathbf{u}_x + \cos^2\theta\sin^2\phi\mathbf{u}_y\mathbf{u}_y + \sin^2\theta\mathbf{u}_z\right) \end{aligned} \tag{A.6}$$

and:

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \mathbf{u}_\phi \mathbf{u}_\phi = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \left(\sin^2\phi \mathbf{u}_x \mathbf{u}_x + \cos^2\phi \mathbf{u}_y \mathbf{u}_y\right) \tag{A.7}$$

Now using:

$$\int_0^{\pi} d\theta \sin\theta = 2 \tag{A.8}$$

$$\int_0^{\pi} d\theta \cos^2\theta \sin\theta = \frac{2}{3} \tag{A.9}$$

$$\int_0^{\pi} d\theta \sin^3\theta = \frac{4}{3} \tag{A.10}$$

$$\int_0^{2\pi} d\phi \sin^2\phi = \int_0^{2\pi} d\phi \cos^2\phi = \pi \tag{A.11}$$

you get

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \mathbf{u}_R \mathbf{u}_R = \frac{4\pi}{3}\hat{\mathbf{1}} \tag{A.12}$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \mathbf{u}_\theta \mathbf{u}_\theta = \frac{2}{3}\pi(\mathbf{u}_x \mathbf{u}_x + \mathbf{u}_y \mathbf{u}_y) + \frac{8}{3}\pi \mathbf{u}_z \mathbf{u}_z \tag{A.13}$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \mathbf{u}_\phi \mathbf{u}_\phi = 2\pi(\mathbf{u}_x \mathbf{u}_x + \mathbf{u}_y \mathbf{u}_y) \tag{A.14}$$

summing up the last two:

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \mathbf{u}_R \mathbf{u}_R = \frac{4}{3}\hat{\mathbf{1}} \tag{A.15}$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta (\mathbf{u}_\theta \mathbf{u}_\theta + \mathbf{u}_\phi \mathbf{u}_\phi) = \frac{8}{3}\hat{\mathbf{1}} \tag{A.16}$$

A.2 Bessel's Functions Indentities

$$\frac{1}{\pi}\int_0^{\pi} d\phi\, e^{i\xi\cos\phi} = J_0(\xi) \tag{A.17}$$

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi\, e^{i\xi\cos\phi} = J_0(\xi) \tag{A.18}$$

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi\, e^{i\xi\cos\phi}\,\cos\phi = iJ_1(\xi) \tag{A.19}$$

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi\, e^{i\xi\cos\phi}\,\cos^2\phi \;=\; J_1'(\xi) = \frac{1}{2}(J_0(\xi) - J_2(\xi)) \tag{A.20}$$

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi\, e^{i\xi\cos\phi}\,\sin^2\phi \;=\; \frac{1}{2}(J_0(\xi) + J_2(\xi)) \tag{A.21}$$

List of publications

1 *Surface plasmon scattering by shallow and deep surface defects*
G. Brucoli and L. Martín-Moreno, arXiv:1009.4137, submitted to Phys.Rev.B.

2 *Comparative study of surface plasmon scattering by shallow ridges and grooves*
G. Brucoli and L. Martín-Moreno, arXiv:1009.1734, submitted to Phys.Rev.B.

3 *Efficient unidirectional ridge excitation of surface plasmons*
I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martín-Moreno, F.J. García-Vidal y and A. Boltasseva, Opt. Express B **17**(9), 7228 (2009).
Vol. 4, Iss. 6 of Virtual Journal for Biomedical Optics.

4 *Theory on the scattering of light and surface plasmon polaritons by arrays of holes and dimples in a metal film*
F de León-Pérez, G. Brucoli, L. Martín-Moreno, F.J. García-Vidal, New Journal of Physics **10**, 105017 (2008).

5 *Efficiency of local surface plasmon polariton excitation on ridges*
I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martín-Moreno, F.J. García-Vidal and A. Boltasseva, Phys. Rev. B **78**, 115115 (2008).

6 *Scattering of surface plasmon polaritons by impedance barriers: Dependence on angle of incidence*
A. Yu. Nikitin, G. Brucoli, L. Martín-Moreno, F.J. García-Vidal, Phys. Rev. B **77**, 195441 (2008).

7 *Surface plasmon polariton scattering by finite-size nanoparticles*
A. B. Evlyukhin, G. Brucoli, L. Martín-Moreno, S.I. Bozhevolnyi and F.J. García-Vidal, Phys. Rev. B **76**, 075426 (2007).

Bibliography

[1] Arfken and Weber, *Mathematical Methods for Physicists* (Harcourt Academic Press, International Edition, 2001).

[2] C.-T.Tai, *Dyadic Green Functions in Electromagnetic Theory* (IEEE Press, New York, 1994).

[3] H.W.Hohmann, Geophysics **40**, 309 (1975).

[4] L. Novotny, B. Hecht, and D. Pohl, J. Appl.Phys **81**, 1798 (1997).

[5] G. Protásio, D. Rogers, and A. Giarola, Radio Sci **17**, 503 (1982).

[6] O. Keller, Phys. Rev. B **34**, 3883 (1986).

[7] L. W. Li, J. Bennet, and P. Dyson, Int.J.Electron. **70**, 803 (1991).

[8] O. J. F. Martin and N. B. Piller, Phys. Rev. E **58**, 3909 (1998).

[9] M. Paulus, P. Gay-Balmaz, and O. J. F. Martin, Phys. Rev. E **62**, 5797 (2000).

[10] L.Novotny and B.Hecht, *Principles of Nano-Optics* (Cambridge University Press, Cambridge, 2006).

[11] J. Wang, *Generalized Moment Methods in Electromagnetics* (Wiley, Chichester UK, 1991).

[12] O. J. F. Martin, C. Girard, and A. Dereux, Phys. Rev. Lett. **74**, 526 (1995).

[13] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (IEEE Press, New York, 2003).

[14] L. Novotny, J. Opt. Soc. Am. A **14(1)**, 105 (1997).

[15] A.D.Yaghjian, Proc.IEE **68**, 248 (1980).

[16] W.C.Chew, IEEE Trans. Antennas Propagat. **37(10)**, 1322 (1989).

[17] W.C.Chew, IEEE Trans. Antennas Propagat. **36(11)**, 0018 (1998).

[18] K. Aki and P. Richard, *Quantitative Seismology-Theory and Methods, vols. I and II* (Freeman, New York, 1980).

[19] A. Sommerfeld, *Partial Differential Equations in Physics* (Academic Press, New York, 1964).

[20] P.M.Morse and H.Feshbach, *Methods of Theoretical Physics, Part 2* (McGraw-Hill, New York, 1953).

[21] J. Bladel, *Electromagnetic Fields* (McGraw-Hill, New York, 1964).

[22] K.-M. Chen, IEEE Proceedings. **65**, 1202 (1977).

[23] C.Gao and T. H. C.Torres-Verdín, Progress In Electromag. Research, PIER. **52**, 47 (2005).

[24] K.-H. D. L. Tsang, J. A. Kong, *Scattering of electromagnetic waves, Vol. 1: Theories and Applications* (Wiley, 2000).

[25] C.J.Bouwkamp, Reports on Progress in Physics **17**, 35 (1954).

[26] O.D.Kellog, *Foundations of Potential Theory(Chapter VI)* (Springer-Verlag, Berlin Heidelberg New York, 1967).

[27] J. Bladel, IEEE Trans. Antennas Propagat. (1961).

[28] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941).

[29] H.G.Fikioris, Journal of Mathematical Physics **Vol.6, No 11**, 1617 (1965).

[30] B.T.Draine and P.J.Flatau, J. Opt. Soc. Am. A **11**, 1491 (1994).

[31] T. Søndergaard and S. I. Bozhevolnyi, Phys. Rev. B **67**, 165405 (2003).

[32] T. Søndergaard and S. I. Bozhevolnyi, Phys. Rev. B **69**, 045422 (2004).

[33] L. . G. F. for Planarly Layered Media, This paper is part of the course: Advanced Electromagnetism (MIT Spring 2003).

[34] A. W.L.Barnes and T. Ebbesen, Nature **424**, 824 (2003).

[35] H. A. Yousif and R. Melka, Computer Physics Communications **106**, 199 (1997).

[36] A. Y. Nikitin, G. Brucoli, F. J. García-Vidal, and L. Martín-Moreno, Phys. Rev. B **77**, 195441 (2008).

[37] A. B. Evlyukhin and S. I. Bozhevolnyi, Phys. Rev. B **71**, 134304 (2005).

[38] A. B. Evlyukhin, G. Brucoli, L. Martín-Moreno, S. I. Bozhevolnyi, and F. J. García-Vidal, Phys. Rev. B **76**, 075426 (2007).

[39] R. W. Keyes, Proc. of IEEE **89**, 227 (2001).

[40] R. Zia, J. A. Schuller, A. Chandran, and M. L. Brongersma, Materials Today **9**, 20 (2006).

[41] W. L. Barnes, A. Dereux, and T. W. Ebbesen, Nature (London) **424**, 824 (2003).

[42] S. Lal, S. Link, and N. J. Halas, Nature Photonics **1**, 641 (2007).

[43] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, Berlin, 1988).

[44] A. V. Zayats and I. I. Smolyaninov, J. Opt. A,Pure Appl. Opt. **5**, S16 (2003).

[45] B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye, and D. W. Pohl, Phys. Rev. Lett. **77**, 1889 (1996).

[46] H. Kano, S. Mizuguchi, and S. Kawata, J. Opt. Soc. Am. B **15**, 1381 (1998).

[47] H. Kano, D. Nomura, and H. Shibuya, Appl. Opt. **43**, 2409 (2004).

[48] Q. Zhan, Opt. Lett. **31**, 1726 (2006).

[49] L. Salomon, G. Bassou, H. Aourag, J. P. Dufour, F. de Fornel, F. Carcenac, and A. V. Zayats, Phys. Rev. B **65**, 125409 (2002).

[50] E. Devaux, T. W. Ebbesen, J.-C. Weeber, and A. Dereux, Appl. Phys. Lett. **83**, 4936 (2003).

[51] F. López-Tejeira, S. G. Rodrigo, L. Martín-Moreno, F. J. García-Vidal, E. Devaux, J. Dintinger, T. W. Ebbesen, J. R. Krenn, I. P. Radko, S. I. Bozhevolnyi, et al., New J. Phys. **10**, 033035 (2008).

[52] H. Ditlbacher, J. R. Krenn, N. Felidj, B. Lamprecht, G. Schider, M. Salerno, A. Leitner, and F. R. Aussenegg, Appl. Phys. Lett. **80**, 404 (2002).

[53] H. Ditlbacher, J. R. Krenn, A. Hohenau, A. Leitner, and F. R. Aussenegg, Appl. Phys. Lett. **83**, 3665 (2003).

[54] A.-L. Baudrion, F. de León-Pérez, O. Mahboub, A. Hohenau, H. Ditlbacher, F. J. García-Vidal, J. Dintinger, T. W. Ebbesen, L. Martín-Moreno, and J. R. Krenn, Opt. Express **16**, 3420 (2008).

[55] G. Lévêque and O. J. F. Martin, J. Appl. Phys. **100**, 124301 (2006).

[56] I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martín-Moreno, F. J. García-Vidal, and A. Boltasseva, Phys. Rev. B **78**, 115115 (2008).

[57] M. Mansuripur, J. Opt. Soc. Am. A **3**, 2086 (1986).

[58] J. D. Jackson, *Classical Electrodynamics, 3rd ed.* (Wiley, New York, 1998).

[59] L. Scaffardi, N. Pellegri, O. de Sanctis, and J. Tocho, Nanotechnology **16**, 158 (2005).

[60] P. B. Johnson and R. W. Christy, Phys.Rev.B **6**, 4370 (1972).

[61] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).

[62] J. Bahns, A. Imre, V. Vlasko-Vlasov, J. Pearson, J. Hiller, L. Chen, and Welp.U, Apl.Phys.Lett. **91**, 081104 (2007).

[63] J. Vuckovic, M. Loncar, and A. Scherer, IEEE J. of Quant. Electr. **36**, 1131 (2000).

[64] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-J. Laluet, and T. W. Ebbesen, Nature (London) **440**, 508 (2006).

[65] A. D. B. L. Y. L. J.-C. Weeber, J. R. Krenn and J. P. Goudonnet, Phys. Rev. B **64**, 045411 (2001).

[66] W. L. Barnes, A. Dereux, and T. W. Ebbesen, Nature (London) **424**, 824 (2003).

[67] S. A. Maier, Current Nanoscience **1**, 17 (2005).

[68] F.López-Tejeira, F.J.García-Vidal and L. Martín-Moreno, Phys. Rev. B **72**, 161405 (2005).

[69] M. U. González, J.-C. Weeber, A.-L. Baudrion, A. Dereux, A. L. Stepanov, J. R. Krenn, E. Devaux, and T. W. Ebbesen, Phys. Rev. B **73**, 155416 (2006).

[70] A. Y. Nikitin, F. López-Tejeira, and L. Martín-Moreno, Phys. Rev. B **75**, 035129 (2007).

[71] A. Drezet, A. Hohenau, A. L. Stepanov, H. Ditlbacher, B. Steinberger, N. Galler, F. R. Aussenegg, A. Leitner, and J. R. Krenn, Appl. Phys. Lett. **89**, 091117 (2006).

[72] R. Zia, J. Sculler, A. Chandran, and M. Brongersman, Materials Today, **9**, 20 (2006).

[73] T. Ebbessen, C. Genet, and S. Bozhevolny, Physics Today **61(5)**, 44 (2008).

[74] S.A.Maier, *Plasmonics: Fundamental and Applications* (Springer-Verlag, New York, 2006).

[75] E. Ozbay, Science **311**, 189 (2006).

[76] J. Krenn, H. Ditlbacher, G. Schider, A. Hohenau, A. Leitner, and F. R. Aussenegg, J. Microsc. **209**, 167 (2003).

[77] J.-C. Weeber, Y. Lacroute, A. Dereux, T. Ebbesen, C. Girard, M. González, and A. Baudrion, Phys. Rev. B **70**, 235406 (2004).

[78] I. P. Radko, S. I. Bozhevolnyi, G. Brucoli, L. Martin-Moreno, F. J. Garcia-Vidal, and A. Boltasseva, Opt. Express **17**, 7228 (2009).

[79] A. V. Zayats, I. Smolyaninov, and A. Maradudin, Physics Reports **408**, 131 (2005).

[80] P. J. Valle, F. Moreno, J. M. Saiz, and F. González, Phys. Rev. B **51**, 13681 (1995).

[81] J. A. Sánchez-Gil, Appl. Phys. Lett., **73**, 3509 (1998).

[82] J. A. Sánchez-Gil and A. A. Maradudin, Phys. Rev. B **60**, 8359 (1999).

[83] F. Pincemin, A. A. Maradudin, A. D. Boardman, and J.-J. Greffet, Phys. Rev. B **50**, 15261 (1994).

[84] J.A.Sanchez-Gil and A.A Maradudin, Appl.Phys.Lett. **86**, 251106 (2005).

[85] J.A.Sanchez-Gil and A.A Maradudin, Opt.Express **12(5)**, 883 (2004).

[86] I. Chremmos, J. Opt. Soc. Am. A **27**, 85 (2010).

[87] G. Lévêque, O. J. F. Martin, and J. Weiner, Phys. Rev. B **76**, 155418 (2007).

[88] M. Paulus and O. Martin, Phys. Rev. E **63**, 066615 (2001).

[89] H.A.Bethe, Phys.Review **66(7)**, 163 (1944).

[90] F.Pincemin and J.J.Greffet, J.Opt.Soc.Am.B **13**, No.7, 1499 (1996).

[91] F.J.García-Vidal, H.J.Lezec, T.W.Ebbessen, L.Martín-Moreno, Phys. Rev. Lett. **90**, 213901 (2003).

[92] T. Søndergaard, S. I. Bozhevolnyi, and A. Boltasseva, Phys. Rev. B **73**, 045320 (2006).

[93] A. Y. Nikitin, G. Brucoli, F. J. García-Vidal, and L. Martín-Moreno, Phys. Rev. B **77**, 195441 (2008).

[94] M. Cottam and D.R.Tilley, *Introduction to Surface and Superlattice Excitations* (Cambridge University Press, Cambridge , U.K., 1989).

[95] M.Kuttge, F.J.García de Abajo and A.Polman, Opt.Express **17(12)**, 10385 (2009).

[96] J.S.Q.Liu, J.S.White, S.Fan and M.L.Brongersma, Opt.Express **17(20)**, 17837 (2009).

[97] F.López-Tejeira, F.J.García-Vidal and L. Martín-Moreno, Appl.Phys.A **89**, 251 (2007).

[98] W. L. Barnes, T. W. Preist, S. C. Kitson, and J. R. Sambles, Phys. Rev. B **54**, 6227 (1996).

[99] S. C. Kitson, W. L. Barnes, and J. R. Sambles, Phys.Rev.Lett **77**, 2670 (1996).

[100] E. Hendry, F. J. Garcia-Vidal, L. Martin-Moreno, J. G. Rivas, M. Bonn, A. P. Hibbins, and M. J. Lockyear, Phys. Rev. Lett. **100**, 123901 (2008).

[101] J.-C. Weeber, A. Dereux, C. Girard, J. R. Krenn, and J.-P. Goudonnet, Phys. Rev. B **60**, 9061 (1999).

[102] P. Berini, Phys. Rev. B **61**, 10484 (2000).

[103] B. Lamprecht, J. R. Krenn, G. Schider, H. Ditlbacher, M. Salerno, N. Felidj, A. Leitner, F. R. Aussenegg, and J. C. Weeber, Appl. Phys. Lett. **79**, 51 (2001).

[104] J.-C. Weeber, J. R. Krenn, A. Dereux, B. Lamprecht, Y. Lacroute, and J. P. Goudonnet, Phys. Rev. B **64**, 045411 (2001).

[105] J.-C. Weeber, Y. Lacroute, and A. Dereux, Phys. Rev. B **68**, 115401 (2003).

[106] J. A. Sánchez-Gil and A. A. Maradudin, Phys. Rev. B **60**, 8359 (1999).

[107] J.-C. Weeber, Y. Lacroute, A. Dereux, E. Devaux, T. Ebbesen, C. Girard, M. U. González, and A.-L. Baudrion, Phys. Rev. B **70**, 235406 (2004).

[108] J. Gómez Rivas, M. Kuttge, H. Kurtz, P. Haring Bolivar, and J. A. Sánchez-Gil, Appl. Phys. Lett. **88**, 082106 (2006).

[109] Nikitin, A. Yu. and Martín-Moreno, L., Phys. Rev. B **75**, 081405 (2007).

[110] F. Lopez-Tejeira, S. G. Rodrigo, L. Martin-Moreno, F. J. Garcia-Vidal, E. Devaux, T. W. Ebbesen, J. R. Krenn, I. P. Radko, S. I. Bozhevolnyi, M. U. Gonzalez, et al., Nat Phys **3**, 324 (2007).

[111] I. I. Smolyaninov, D. L. Mazzoni, J. Mait, and C. C. Davis, Phys. Rev. B **56**, 1601 (1997).

[112] C. Girard and A. Dereux, Phys. Rev. B **49**, 11344 (1994).

[113] A. B. Evlyukhin and S. I. Bozhevolnyi, Phys. Rev. B **71**, 134304 (2005).

[114] S. I. Bozhevolnyi and V. Coello, Phys. Rev. B **58**, 10899 (1998).

[115] I. P. Radko, S. I. Bozhevolnyi, A. B. Evlyukhin, and A. Boltasseva, Opt. Express **15**, 6576 (2007).

[116] M. I. Tribelsky and B. S. Luk'yanchuk, Phys. Rev. Lett. **97**, 263902 (2006).

[117] Z. B. Wang, B. S. Luk'yanchuk, M. H. Hong, Y. Lin, and T. C. Chong, Phys. Rev. B **70**, 035418 (2004).

[118] G. Lévêque and O. J. Martin, Opt. Lett. **31**, 2750 (2006).

[119] G. Lévêque and O. J. F. Martin, Opt. Express **14**, 9971 (2006).

[120] U. Hohenester and J. Krenn, Phys. Rev. B **72**, 195429 (2005).

[121] R. Fuchs, Phys. Rev. B **11**, 1732 (1975).

[122] C. Noguez, J.Phys. Chem. C **111**, 3806 (2007).

[123] C.F.Bohren and D.R.Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).